



UNIVERSIDAD POLITÉCNICA DE MADRID



ESCUELA UNIVERSITARIA DE INGENIERÍA TÉCNICA DE TELECOMUNICACIÓN

Ingeniería Técnica de Telecomunicación, especialidad Sonido e Imagen

Curso Académico 2012/2013

Proyecto Fin de Carrera

DOCUMENTACIÓN DE ESTUDIO PARA SISTEMAS AUDIOVISUALES

Autor: Jairo David Guerrero Vásquez

Tutora: Elena Blanco Martín



PROYECTO FIN DE CARRERA PLAN 2000

E.U.I.T. TELECOMUNICACIÓN

TEMA: Sistemas Audiovisuales

TÍTULO: Documentación de estudio para Sistemas Audiovisuales

AUTOR: Jairo David Guerrero Vásquez

TUTOR: Elena Blanco Martín

Vº Bº.

DEPARTAMENTO: DIAC

Miembros del Tribunal Calificador:

PRESIDENTE: Margarita Millán Valenzuela

VOCAL:

VOCAL SECRETARIO: Jose Luis Rodríguez Vázquez

DIRECTOR: Elena Blanco Martín

Fecha de lectura: 20/09/2013

Calificación:

El Secretario,

RESUMEN DEL PROYECTO:

Este proyecto es una documentación sintetizada, para los alumnos de Grado en Imagen y Sonido, de todos los conceptos que conciernen a la asignatura Sistemas Audiovisuales.

La documentación se complementa al final de cada unidad didáctica con un test de evaluación que a su vez ha sido publicado dentro del entorno de Moodle. Este documento proporciona los suficientes recursos de estudio para la asignatura anteriormente citada. Los conceptos clave de cada unidad didáctica se han utilizado para la elaboración de los cuestionarios, de tal manera que los alumnos puedan saber si ha comprendido lo que se explica en la documentación.

La unificación de materia para la asignatura Sistemas Audiovisuales se basa en la recopilación de libros de diversos autores, páginas web y catálogos de productos de empresas del sector audiovisual.

El documento se ha dividido en dos bloques temáticos correspondientes a los temas:

1- Dispositivos de captación y reproducción de sonido e imagen.

2- Señales y formatos de audio y vídeo.

Cada bloque a su vez contiene las unidades didácticas con el conjunto toda la materia de dicho bloque.

Cada unidad se ha desarrollado de manera independiente de las demás pero con relación entre sí, es decir, que cada unidad puede ser estudiada de manera independiente sin que le hagan falta la documentación de otras unidades didácticas para comprender la/s que interesa/n.

RESUMEN

Este proyecto es una documentación sintetizada, para los alumnos de Grado en Imagen y Sonido, de todos los conceptos que conciernen a la asignatura Sistemas Audiovisuales. No obstante puede servir para todo aquel al que le interese la materia, sin ser necesariamente estudiante.

El material se basa en la recopilación de libros de diversos autores, páginas web y catálogos de productos de empresas del sector audiovisual. Se intenta con esto incentivar en el auto-aprendizaje, proporcionando multitud de fuentes de información.

El documento se ha dividido en dos bloques temáticos correspondientes a los temas:

- 1- Dispositivos de captación y reproducción de sonido e imagen.
- 2- Señales y formatos de audio y vídeo.

Aunque no es tema de este proyecto pero si de la asignatura hay que nombrar el tercer bloque temático, Introducción a los sistemas de transmisión de audio y vídeo. Dado que hay suficiente documentación de estudio sobre éste se ha optado por no incluirlo.

Cada bloque temático a su vez contiene cuatro unidades didácticas. Cada unidad se ha desarrollado de manera independiente a las demás, es decir, que cada unidad puede ser estudiada sin necesidad de recurrir a otras unidades para comprender la/s que interesa/n. Por otro lado hay que remarcar que todos los capítulos tienen relación entre sí.

La documentación se complementa al final de cada unidad didáctica con un test de evaluación que a su vez ha sido publicado dentro del entorno de Moodle en la página correspondiente a la asignatura. Para ello se ha accedido a esta plataforma *on line* con el rol de editor de contenido.

Para la elaboración de los cuestionarios se han tomado los conceptos clave de cada unidad didáctica, de esta manera los alumnos pueden saber si han comprendido lo que se explica en la documentación y mejorar así sus conocimientos.

Para la redacción y estructuración de cada unidad didáctica, así como el documento en general, se ha cogido como referencia la Taxonomía de objetivos de la Educación o Taxonomía de Bloom. Dado que el dominio cognitivo del lector se encuadra dentro del ‘nivel de comprensión’, el documento no resulta tedioso en su estudio. No obstante introduce al alumno en los temas más importantes de la materia, proporcionando una base sólida de conocimiento en sistemas audiovisuales.

Es precisamente el interés en hacer lo más accesible posible este documento lo que ha dificultado su elaboración, ya que el área de estudio es muy extensa y es difícil sintetizar sin eliminar contenido importante. No obstante para hacer más fiable el documento se ha seguido las pautas temáticas y argumentales marcadas por el Departamento Ingeniería Audiovisual y Comunicaciones de la Escuela Universitaria de Ingeniería Técnica de Telecomunicaciones de la Universidad Politécnica de Madrid verificando cada uno de los capítulos con los profesores de este departamento.

Al tratarse de un proyecto con fines académicos, el texto se ha apoyado por figuras, esquemas, tablas, anexos y desarrollo de ecuaciones para hacer más comprensible lo que se expone. Algunos de estas informaciones se incluyen en inglés y no se ha creído conveniente su traducción dado que gran parte de la información que encontrará el alumno a lo largo de la carrera vendrá escrita en este idioma.

Por último hay que decir al lector que es conveniente, pero no necesario, tener ciertas nociones de cálculo, álgebra, ondas y circuitos para seguir con fluidez lo que a continuación se expone.

ABSTRACT

This project concerns all the concepts and topics of the subject Audiovisual Systems.

It has been created for students of Sound and Image Degree, however everyone who's interested in this subject could use it even if isn't a student.

The document is divided into two main thematic sections corresponding to the topics:

- 1- Catchment and reproduction devices of sound and image.
- 2- Audio and video signals and formats.

Even if this subject it isn't mention in this project, it's very important to quote a third important thematic of this block, such as Introduction about Transmission of Audio and Video System. Since there is enough study-documentation about this topic, it has been taken the choice to don't integrate this chapter in this project.

Every thematic block in this project is divided in chapters that have been developed in an independent way: that's means that for each unit it is not necessary to look forward to other chapters in this project. On the other hand it is necessary to emphasize that all the chapters are related one to each other.

Every didactic unit and chapter ends with an evaluation test, that has been published with Moodle System using a content editor account. Those exercises will help in a easy way the student to improve his skills and his own ability.

Collection of books of various authors, websites and product catalogs of audiovisual companies are used in this document and are included for stimulate the curiosity of the student.

The key concepts of each unit have been used for making tests, so in this way students could be able to know if they have understood what the documentation explains and improve his skills.

For writing and building each didactic unit, such as in the general document, it has been taken reference from Bloom's Taxonomy. Since the skills and competence of the student are concentrated in the 'comprehension level', it will not be very complicated or hard to study. In spite of everything, all of thematic treated and discussed in documentation gives a solid knowledge of topic about audiovisual systems. The most difficult thing of realizing this document it was to take very complex topic and try to explain them as simply as possible.

In spite of everything for making this document as much accurated as possible it has been taken as point of reference rules established by the Department of Audiovisual Engineering & Communications of University School of Telecommunications Engineering (EUITT-UPM).

This Project reach academic goal, for this reason in this document images, tables, annexes and outlines are enclosed in this document for an easier compression.

At last, it's necessary to say that each lector must have necessary a basic knowledge about arithmetic, calculus, waves and electronic circuits in the order that he could follow in a fluently way what the documentation set out..

*A mis PADRES,
a mis HERMANAS,
a UZIEL.*

A todo aquel que me ayudo a estar aquí y al que me lo puso difícil.

Gracias.

ÍNDICE DE CONTENIDOS

INTRODUCCIÓN	3
TEMA 1	5
1 MICRÓFONOS	7
1.1 Introducción a la señal de audio	7
1.2 Definición y características de los micrófonos.....	9
1.3 Tipos de micrófonos.....	14
1.3.1 Según su TAM.....	14
1.3.2 Según su TME	17
1.4 Conexionado	26
1.4.1 Tipos de conexiones	27
1.5 Cuestionario.....	30
2 ALTAVOCES	33
2.1 Definición de altavoz.....	35
2.2 Clasificación de los altavoces	36
2.2.1 Según su TEM	36
2.2.2 Según la banda de frecuencias reproducida	40
2.2.3 Según el TMA	41
2.3 Altavoz dinámico de bobina móvil.....	41
2.3.1 Características constructivas de un altavoz de bobina móvil.....	41
2.4 Especificaciones de un altavoz.....	49
2.4.1 Impedancia	50
2.4.2 Potencia.....	51
2.4.3 Sensibilidad	52
2.4.4 Respuesta en frecuencia	53
2.4.5 Directividad	54
2.4.6 Eficiencia	56
2.4.7 Distorsión	57
2.4.8 Parámetros de Thiele-Small.	58
2.5 Cajas acústicas	58
2.5.1 Conexión de las cajas acústicas.....	59
2.6 Cuestionario.....	61
3 CÁMARAS Y SENSORES	65
3.1 Introducción a la señal de vídeo	65

3.1.1	Mezcla aditiva espacial. Agudeza visual	66
3.1.2	Mezcla aditiva temporal. Memoria visual	67
3.2	Sistema de transmisión de imágenes	68
3.2.1	Número de líneas	69
3.2.2	Ancho de banda.....	69
3.2.3	Resolución	70
3.2.4	Sistema de exploración	70
3.2.5	Corrección gamma.....	73
3.2.6	Sistemas de vídeo en color	73
3.3	Cámaras de Vídeo	76
3.3.1	Diagrama general de una cámara.....	77
3.3.2	Bloque Óptico.....	78
3.3.3	Balance de blancos	79
3.3.4	Cuantificación de la Iluminación.....	79
3.3.5	Sensibilidad de una Cámara	80
3.3.6	Lentes	80
3.3.7	Tamaño del sensor y de la imagen	81
3.3.8	Distancia focal	81
3.3.9	Apertura del diafragma	81
3.3.10	Profundidad de campo	82
3.3.11	MTF.....	82
3.3.12	Sensores	83
3.4	Tratamiento de las señales RGB.....	89
3.5	Cámaras analógicas y digitales	90
3.6	Cámaras con soporte de grabación: Camcorders	90
3.7	Funcionamiento de una cámara de vídeo	90
3.8	Cuestionario	92
4	MONITORES Y PROYECTORES	95
4.1	Conceptos básicos.....	95
4.2	Conectores	98
4.3	Tecnologías usadas en Monitores.....	102
4.3.1	Monitores CRT (Tubo de rayos catódicos).....	102
4.3.2	Monitores LCD	103
4.3.3	Monitores LED	111

4.3.4	Monitores plasma o PDP	111
4.3.5	PLED (Pixel Light Emitting Display).....	114
4.3.6	Tecnología de visualización OLED	114
4.3.7	Tecnología de visualización SED.....	118
4.3.8	Tecnología láser	120
4.3.9	Tecnología DLP	122
4.3.10	Tecnología LCoS.....	124
4.4	Cuestionario.....	128
TEMA 2		133
5	DIGITALIZACIÓN DE LAS SEÑALES DE AUDIO Y VÍDEO.....	135
5.1	Muestreo.....	135
5.2	Cuantificación y codificación.....	135
5.3	Digitalización de audio	136
5.3.1	Muestreo de la señal de audio	136
5.3.2	Cuantificación de la señal de audio.....	137
5.3.3	Señal digital de audio AES3	139
5.4	Digitalización de vídeo	141
5.4.1	Señales digitalizadas.....	142
5.4.2	Muestreo.....	143
5.4.3	Cuantificación.....	146
5.4.4	Señal de Digital de Vídeo: SDI	147
5.5	Conversión digital-analógica	149
5.6	Cuestionario.....	151
6	CODIFICACIÓN DE LA SEÑAL DE AUDIO.....	155
6.1	Psicoacústica	155
6.2	Fisiología del oído humano y bandas críticas. Percepción auditiva	157
6.2.1	Umbral de audición	157
6.2.2	Enmascaramiento frecuencial.....	157
6.2.3	Enmascaramiento temporal.....	159
6.3	Fundamentos de la codificación perceptual	160
6.3.1	Asignación de bits	161
6.4	Métodos de codificación.....	162
6.4.1	Codificación digital de la forma de onda.....	162
6.4.2	Codificación de audio	163

6.5	Estándares de codificación.....	170
6.5.1	Estándar de audio MPEG-1.....	170
6.5.2	Estándar de audio MPEG-2.....	173
6.5.3	MPEG-2 AAC.....	175
6.5.4	Codificador AC-3 (DOBLY DIGITAL).....	177
6.6	Cuestionario.....	180
7	CODIFICACIÓN DE LA SEÑAL DE VÍDEO.....	183
7.1	Principios de compresión.....	184
7.1.1	Compresión con pérdida y sin pérdida.....	184
7.1.2	Redundancia.....	185
a)	Redundancia espacial.....	185
b)	Redundancia temporal.....	185
c)	Redundancia de codificación.....	185
d)	Redundancia Psicovisual.....	186
7.2	Compresión y codificación MPEG.....	186
7.2.1	Descomposición de la imagen.....	186
7.2.2	Concepto macrobloque.....	187
7.2.3	Señales utilizadas.....	187
7.2.4	Transformada discreta del coseno.....	189
7.2.5	Cuantificación.....	190
7.2.6	Codificación tipo Intra.....	191
7.2.7	Codificación tipo Inter.....	194
7.2.8	Estructura jerárquica de la señal.....	202
7.2.9	Orden de procesamiento de las imágenes.....	203
7.3	Cuestionario.....	204
8	SOPORTES DE ALMACENAMIENTO DE AUDIO Y VÍDEO.....	207
8.1	Memorias en estado sólido.....	208
8.1.1	Memorias internas.....	208
8.1.2	Tarjetas compact flash CF.....	209
8.1.3	Tarjeta SD (Secure Disk).....	210
8.1.4	Tarjeta Memory Stick.....	210
8.1.5	Tarjeta XD.....	211
8.2	Soportes ópticos.....	211
8.2.1	Disco compacto o CD (Compact Disk).....	212

8.2.2	DVD	214
8.2.3	Blu-ray	217
8.3	Soportes magneto-ópticos.....	221
8.3.1	MiniDisc.....	221
8.4	Soportes Magnéticos	221
8.4.1	Cinta magnética.....	222
8.4.2	Discos Duros.....	225
8.5	Formatos de archivado	230
8.5.1	Códec.....	230
8.5.2	Formatos de audio	231
8.5.3	Formatos de vídeo.....	235
8.5.4	Formatos de archivado de imagen.....	238
8.6	Cuestionario	242
9	CONCLUSIONES.....	247
10	BIBLIOGRAFÍA	249
ANEXO A	253
	Dithering	253
	Cálculo de la Relación señal/ruido	254
ANEXO B	257
	Obtención de los códigos VLC para los bloques de luma y croma.....	257
	Codificación de los índices AC tipo Intra	258

ÍNDICE DE FIGURAS

Figura 1-1. Representación de la señal de audio.....	7
Figura 1-2. Esquema de un sistema de sonido.	7
Figura 1-4. Símbolo del micrófono. ‘p’-presión incidente, ‘E’-señal eléctrica.....	9
Figura 1-5. Transductores que conforman el micrófono.....	9
Figura 1-6. Respuesta en frecuencia de un micrófono.....	10
Figura 1-7. Coordenadas esféricas para la determinación de la directividad de un micrófono.....	11
Figura 1-8. Diagrama polar de Directividad a diferentes frecuencias	12
Figura 1-9. Condiciones en las que se realiza la medida del diagrama polar de Directividad.....	12
Figura 1-10. Patrones más comunes de directividad.	13
Figura 1-11. Ancho de haz para -3 dB y -6 dB.....	13
Figura 1-12. TAM del micrófono de presión.....	15
Figura 1-13. Cuando la onda incide con $\theta=180^\circ$	15
Figura 1-14. TAM del micrófono de gradiente.	16
Figura 1-15. Efecto proximidad de los micrófonos de gradiente.	16
Figura 1-16. Diagrama polar de directividad en baja frecuencia (izqda.) y en alta frecuencia (dcha.)	17
Figura 1-17. TMA del micrófono combinado de gradiente y presión.....	17
Figura 1-18. Micrófono dinámico de bobina móvil funcionando como micrófono de presión.....	18
Figura 1-19. Esquema de un micrófono dinámico de bobina móvil.....	18
Figura 1-20. Micrófono dinámico de cinta funcionando como micrófono de gradiente de presión.	19
Figura 1-21. Micrófono electroestático de condensador funcionando como micrófono de presión.	20
Figura 1-22. Alimentación <i>Phantom</i> para micrófonos de condensador.....	21
Figura 1-23. Esquema constructivo de un micrófono de electret con diafragma metálico	21
Figura 1-24. Esquema de un micrófono electret alimentado con una pila de 1,5V.	21
Figura 1-25. Construcción de un micrófono de carbón.....	22
Figura 1-26. Construcción de un micrófono piezoeléctrico.....	22
Figura 1-27. Esquema constructivo de un micrófono de doble diafragma.	23
Figura 1-28. Comportamiento de un micrófono de cañón.....	24
Figura 1-29. Micrófono con paraboloide.....	24
Figura 1-30. Captación del sonido de la cápsula microfónica de los micrófonos con paraboloide.....	25
Figura 1-31. Esquema de funcionamiento de los micrófonos inalámbricos.....	25
Figura 1-32. Diversidad espacial activa o ‘verdadera’.	26
Figura 1-33. Antenas receptoras	26

Figura 1-34. Macho (13) y hembra (25) del XLR.	26
Figura 1-35. Estructura del cable coaxial.	27
Figura 1-36. Esquema eléctrico de la conexión asimétrica.....	27
Figura 1-37. Esquema del envío balanceado de señal.....	28
Figura 1-38. Par trenzado apantallado.	28
Figura 1-39. Esquema eléctrico del envío balanceado.	28
Figura 1-40. Rechazo del ruido en conexión balanceada.	29
Figura 2-1. Esquema de un sistema de sonido.	33
Figura 2-2. Mesa de mezclas modelo M7CL de Yamaha.....	33
Figura 2-3. Amplificador de potencia modelo D100 de DAS Audio.	34
Figura 2-4. Etapas de un amplificador de potencia.	34
Figura 2-5. Sistema pasivo de filtrado.	34
Figura 2-6. Sistema activo de filtrado.	35
Figura 2-7. Altavoz.	35
Figura 2-8. Esquema genérico de un altavoz conectado a un amplificador.	35
Figura 2-9. Altavoz de bobina móvil	36
Figura 2-10. Altavoz dinámico de cinta.....	37
Figura 2-11. Transductor isodinámico. Disposición de los imanes, diafragma y conductor.....	37
Figura 2-12. Disposición de la bobina plana.	38
Figura 2-13. Sentido del campo B, la fuerza $f(t)$ y la corriente $i(t)$ sobre pistas contiguas.	38
Figura 2-14. Detalle del altavoz electroestático	38
Figura 2-15. Corte esquemático de un altavoz piezoeléctrico lastrado por un diafragma de papel.	39
Figura 2-16. Configuración de un bimorph.....	39
Figura 2-17. Diseño simplificado del auricular del teléfono de Bell	40
Figura 2-18. Corte esquemático de un altavoz de bobina móvil.	41
Figura 2-19. Corte en perspectiva de un altavoz dinámico de bobina móvil.....	42
Figura 2-20. Geometría de una cúpula convexa	42
Figura 2-21. Detalle de un ecualizador de fase.....	43
Figura 2-22. Respuesta en frecuencia típica de un altavoz de cúpula debida a la pérdida de fase.....	43
Figura 2-23. Diversos perfiles de diafragmas cónicos.....	43
Figura 2-24. Diafragma elíptico.....	44
Figura 2-25. Estructura de los altavoces de doble cono.	44
Figura 2-26. Estructura de un altavoz coaxial.	44

Figura 2-27. Corte transversal del altavoz de bobina móvil.	45
Figura 2-28. Bobina cuadrada y circular dispuestas en la cara externa del former.	45
Figura 2-29. Imán cilíndrico metálico con orificio central de ventilación.....	46
Figura 2-30. Distorsión en baja frecuencia debido a la salida de la bobina del entrehierro	47
Figura 2-31. Bobina mayor a la altura del entrehierro	47
Figura 2-32. Suspensión primaria o araña (izqda.). Doble suspensión en un woofer (dcha.)	48
Figura 2-33. Suspensión secundaria de media vuelta (1),	48
Figura 2-34. Tapadera.....	49
Figura 2-35. Detalle de la carcasa y culata de un altavoz de graves.....	49
Figura 2-36. Bafle estándar recomendado por la norma IEC.	50
Figura 2-37. Impedancia eléctrica de entrada, $ Z_{ee}(f) $ [Ω] , de un altavoz coaxial de doble vía.	50
Figura 2-38. Función de transferencia (gálibo de tolerancia) para el filtrado de ruido según la norma IEC.	51
Figura 2-39. Función de transferencia para el filtrado de ruido según AES.	52
Figura 2-40. Respuesta en frecuencia obtenida de un barrido de tonos.	53
Figura 2-41. Respuesta en frecuencia obtenida de excitación con tercios de octava.....	53
Figura 2-42. Respuesta en frecuencia obtenida de la función de transferencia del altavoz.	54
Figura 2-43. Configuración geométrica para las curvas polares de directividad.....	54
Figura 2-44. Diagramas polares de directividad medidos en 1/3 de octava.	55
Figura 2-45. D_{lax} y Q_{ax} en función de la frecuencia.	55
Figura 2-46. Ángulos de cobertura a -3, -6 y -9 dB para todos los planos del altavoz.....	56
Figura 2-47. Ángulos de cobertura para -3, -6 y -9 dB. El trazo más grueso corresponde a los -6dB.	56
Figura 2-48. Distorsión armónica en función de la frecuencia	58
Figura 2-49. Caja acústica cerrada.....	58
Figura 2-50. Caja acústica abierta con un sistema de Bass-reflex.	59
Figura 2-51. Caja con radiador pasivo.	59
Figura 2-52. Sistema sonoro de un <i>home cinema</i>	60
Figura 3-1. Espectro electromagnético conocido.	65
Figura 3-2. Curvas de respuesta del ojo	65
Figura 3-3. Niveles de brillo. De mayor a menor de izquierda a derecha.	66
Figura 3-4. Efecto de la variación en el brillo.	66
Figura 3-5. Efecto de la variación en el contraste.	66
Figura 3-6. Captación de dos puntos diferenciados.	67
Figura 3-7. Captación de dos puntos diferenciados como uno solo.....	67

Figura 3-8. Posiciones relativas de los luminóforos en las pantallas de televisión.....	67
Figura 3-9. Respuesta del ojo a un impulso.	68
Figura 3-10. Respuesta del ojo a varios impulsos seguidos.	68
Figura 3-11.Cálculo del número de líneas.	69
Figura 3-12. Señal transmitida para calcular la máxima frecuencia de variación.....	70
Figura 3-13. Sistema de exploración progresivo.....	71
Figura 3-14. Ley de Ferry Porter.	72
Figura 3-15. Escaneado entrelazado.....	72
Figura 3-16. Curvas de corrección Gamma.	73
Figura 3-17. Esquema de la televisión en color.	73
Figura 3-18. Representación de la luminancia, tono y saturación.....	74
Figura 3-19. Funciones de XYZ CIE 1931.....	74
Figura 3-20. Diagrama de cromaticidad con los espacios de color de NTSC, SMPTE y EBU/PAL.....	75
Figura 3-21. Compatibilidad directa.	75
3-22. Compatibilidad inversa.	75
Figura 3-23. Señales RGB y Y Pb Pr de una imagen de barras de color.	76
Figura 3-24. Filtro dicróico.....	77
Figura 3-25. Esquema general de una cámara de vídeo.....	77
Figura 3-26. Distribuciones espectrales de la luz solar y el Tungsteno.....	78
Figura 3-27. Iluminación	79
Figura 3-28. Detalle del bloque óptico de una cámara.....	80
Figura 3-29. Distancia Focal.	81
Figura 3-30. Obturador con diferentes aperturas (distintos números F).	82
Figura 3-31. Diferentes resultados obtenidos del análisis de la MTF.	83
Figura 3-32. Esquema representativo de un array sensor con fotodiodos cuadrados	83
Figura 3-33. Sensor CCD (izqda.) y CMOS (dcha.).	84
Figura 3-34. Arquitectura <i>Frame Transfer</i>	85
Figura 3-35. Arquitectura <i>Interline Transfer</i>	86
Figura 3-36. Arquitectura <i>Frame-Interline Transfer</i>	87
Figura 3-37. Un sensor CMOS.....	88
Figura 3-38. Cámara de vídeo tipo <i>Camcorder</i>	90
Figura 4-1. Monitor profesional Panasonic.	95
Figura 4-2. Proyector de vídeo.....	95

Figura 4-3. Comparación de resoluciones entre los diferentes formatos de vídeo.	97
Figura 4-4. Tamaño del punto.	97
Figura 4-5. Conector RCA para SCVF. Hembra RCA (arriba) y macho RCA (abajo).	98
Figura 4-6. Conector para S-Vídeo: DIN 4 pins.	98
Figura 4-7. Conectores RCA para vídeo en componentes.	99
Figura 4-8. Conector SUB-D15 para VGA.	99
Figura 4-9. Conector para HDMI.	99
Figura 4-10. Pines del HDMI.	100
Figura 4-11. Conector DVI-D usado solo para señales digitales.	100
Figura 4-12. Pineado del DVI dependiendo del tipo de señal que se transmita.	101
Figura 4-13. Interior de una pantalla de televisión.	102
Figura 4-14. Estructura de una molécula LC.	103
Figura 4-15. Fase nemática.	104
Figura 4-16. Onda electromagnética propagándose en la dirección de u . λ es la longitud de onda.	104
Figura 4-17. Polarización vertical.	104
Figura 4-18. Reorientación de las moléculas del cristal líquido nemático.	105
Figura 4-19. Geometría torcida producida al girar las placas de vidrio 90°	106
Figura 4-20. Elementos de una celda de un indicador electro-óptico.	106
Figura 4-21. Imagen típica de un indicador digital electro-óptico.	107
Figura 4-22. Capas de una pantalla LCD.	107
Figura 4-23. Capas de un monitor LCD: modo transmisivo.	108
Figura 4-24. Estructura de una matriz pasiva de LCD.	108
Figura 4-25. Pantalla de matriz pasiva con filtros de colores básicos.	109
Figura 4-26. Estructura de una pantalla de matriz activa.	109
Figura 4-27. Sección de una pantalla LCD TFT.	109
Figura 4-28. Distribución en tiras de los píxeles.	110
Figura 4-29. Distribución en mosaico de los píxeles.	110
Figura 4-30. Distribución en Delta de los píxeles.	110
Figura 4-31. Esquema de proyectores LCD.	111
Figura 4-32. Composición de una pantalla de plasma.	112
Figura 4-33. Principio de funcionamiento del plasma.	112
Figura 4-34. Efecto pantalla quemada.	113
Figura 4-35. Estructura del OLED.	115

Figura 4-36. Capas de emisión y conducción del OLED.	115
Figura 4-37. Principio de funcionamiento del OLED.	116
Figura 4-38. La tecnología OLED permiten elasticidad en los elementos de visualización.....	116
Figura 4-39. Comparativa entre TFT y OLED.	117
Figura 4-40. La diferencia entre SED y CRT	118
Figura 4-41. Estructura de un SED.	118
Figura 4-42. Proceso de iluminación de los píxeles en paneles SED.....	119
Figura 4-43. Comparativa de imágenes en Plasma, SED y LCD.	119
Figura 4-44. Iluminación de los píxeles: emisión un rayo de cada uno de los tres colores primarios.	120
Figura 4-45. Estructura de un píxel con láser Necsel.	121
Figura 4-46. Esquema de proyección con la tecnología Necsel.	121
Figura 4-47. Comparativa Láser (izquierda) vs. Plasma (derecha).	122
Figura 4-48. Tamaño de una matriz de visualización con tecnología DLP.	122
Figura 4-49. Ejemplo de proyección con 4 bits con un proyector DLP.	122
Figura 4-50. Estructura de los espejos reflectantes del DMD.....	123
Figura 4-51. Rueda de color.	123
Figura 4-52. Proceso de proyección con un solo chip.....	123
Figura 4-53. Efecto arcoíris.	124
Figura 4-54. Estructura del LCOS.	125
Figura 4-55. Detalle de la formación de la imagen en el LCOS.	125
Figura 4-56. Proceso de proyección:	126
Figura 5-1. Etapas de la digitalización.....	135
Figura 5-2. Esquema básico de un sistema de digitalización.	136
Figura 5-3. Dominio del tiempo (izqda.) y dominio de la frecuencia (dcha.).....	136
Figura 5-4. Proceso de cuantificación no uniforme	138
Figura 5-5. Representación de la Ley A con diferentes valores para su parámetro A.	139
Figura 5-6. Representación de la Ley μ con diferentes valores para su parámetro μ	139
Figura 5-7. Diagrama de bloques de la digitalización de audio AES/EBU.	140
Figura 5-8. Forma de onda de la señal AES/EBU.	141
Figura 5-9. Muestreo ortogonal estático.....	143
Figura 5-10. Proceso de digitalización de señales	145
Figura 5-11. Filtro anti-aliasing para la señal de luminancia.	145
Figura 5-12. Filtro anti-aliasing para la señal de crominancia.	146

Figura 5-13. Cuantificador-decuantificador.....	146
Figura 5-14. Línea activa digital para sistemas de 625 líneas.....	147
Figura 5-15. Multiplicación de los tres canales de datos.....	148
Figura 5-16. Multiplexación de los tres canales de datos.....	148
Figura 5-17. Esquema que sigue la señal de vídeo hasta convertirse en SDI.	149
Figura 6-1. Las curvas de igual sonoridad Robinson-Dadson	156
Figura 6-2. El umbral de audición describe los sonidos más débiles que son audibles.....	158
Figura 6-3. Los umbrales de enmascaramiento.....	158
Figura 6-4. El enmascaramiento temporal	159
Figura 6-5. Contorno tiempo-frecuencia	160
Figura 6-6. Codificador DCPM	162
Figura 6-7. Decodificación DCPM. $\tilde{e}(n)$: error cuantificado, $x_2(n)$: dato estimado, $x(n)$: secuencia inicial.	162
Figura 6-8. $x(n)$: dato de entrada.....	163
Figura 6-9. $e(n)$: error de predicción.	163
Figura 6-10. Estructura básica de un codificador y un decodificador frecuencia-tiempo.....	164
Figura 6-11. El efecto de los pre-ecos hace aparecer señal antes del impulso de la señal que los origina.....	164
Figura 6-12. Los codificadores sub-banda crean bandas estrechas de muestras de alta resolución	165
Figura 6-13. Esquema de funcionamiento de la codificación por sub-bandas.....	166
Figura 6-14. Un codificador sub-banda	167
Figura 6-15. El algoritmo de asignación fija el número de bits	168
Figura 6-16. Esquema de funcionamiento de la codificación por transformada	169
Figura 6-17. Estructura de tramas en los layers I, II y III del estándar ISO/MPEG-1.....	172
Figura 6-18. El estándar de audio MPEG-2.....	173
Figura 6-19. Codificador y decodificador de audio ISO/MPEG-2.....	174
Figura 6-20. Diagrama de bloques del codificador MPEG-2 AAC.	176
Figura 6-21. Codificador y decodificador de transformada adaptativa AC-3	179
Figura 7-1. Descomposición de la imagen en macrobloques.	187
Figura 7-2. Descomposición en bloques de un macrobloque.	187
Figura 7-3. Muestras de luminancia y crominancia para 4:4:4.	188
Figura 7-4. Muestras de luminancia y crominancia (formato 4:2:2).	188
Figura 7-5. Muestras de luminancia y crominancia (formato 4:2:0).	189
Figura 7-6. Numeración de los bloques dentro del macrobloque.....	189
Figura 7-7. Conjunto de imágenes base.	190

Figura 7-8. Bloque de cuantificación que se está estudiando.	190
Figura 7-9. Valores implicados en la codificación diferencial.	191
Figura 7-10. Codificación de una imagen tipo Intra.....	192
Figura 7-11. Proceso básico para un bloque.....	192
Figura 7-12. Codificación diferencial de los coeficientes DC.	193
Figura 7-13. Dos posibles formas de exploración en zig-zag.	193
Figura 7-14. Obtención de la imagen en el codificador.	194
Figura 7-15. Obtención de las informaciones a enviar por el codificador.	195
Figura 7-16. Vector de movimiento para el balón.	195
Figura 7-17. Codificación de una imagen tipo P.	196
Figura 7-18. Ejemplo de un conjunto de imágenes.	197
Figura 7-19. Codificación tipo P	197
Figura 7-20. Posibilidades de las imágenes tipo B.....	198
Figura 7-21. Predicciones para las imágenes codificadas tipo B.....	199
Figura 7-22. Imágenes utilizadas para obtener las predicciones.	200
Figura 7-23. Establecimiento del ritmo de imágenes I, P y B.....	200
Figura 7-24. Estructura típica de distribución de imágenes.	201
Figura 7-25. Estructura jerárquica de la señal de vídeo digital MPEG.....	202
Figura 7-26. Alteración en el orden de codificación de las imágenes.....	203
Figura 8-1. Memoria en estado sólido de Microsoft Xbox 360.....	209
Figura 8-2. Tarjeta compact flash de SanDisk.....	209
Figura 8-3.Tarjeta SD de SanDisk.....	210
Figura 8-4. <i>Memory Stick</i> ProDuo de Sony.	210
Figura 8-5. Tarjeta XD de Toshiba.....	211
Figura 8-6. Ejemplo de capas del Disco Óptico Compacto (CD).....	211
Figura 8-7. Cada línea horizontal son las pistas de un disco óptico, en este caso el de un CD.....	212
Figura 8-8. Capas de un CD-R.....	214
Figura 8-9. Capas del CD-RW.	214
Figura 8-10. Comparación entre el CD y el DVD.	215
Figura 8-11. MiniDisc de Sony.	221
Figura 8-12. Hay diferentes sistemas de formatos magnéticos como el VHS, el <i>cassette</i> o el disquete.....	222
Figura 8-13. Esquema de la grabación transversal de la cinta magnética.	223
Figura 8-14. Esquema de grabación lineal en cinta magnética.	223

Figura 8-15. Esquema de grabación helicoidal en cinta magnética.	224
Figura 8-16. En el almacenamiento en cinta magnética	225
Figura 8-17. Disco duro.....	225
Figura 8-18. Estructura interna de un disco duro.....	226
Figura 8-19. Esquema de funcionamiento del RAID-0.....	228
Figura 8-20. Esquema de funcionamiento del RAID-1.....	229
Figura 8-21. Esquema de funcionamiento del RAID-3.....	229
Figura 8-22. Esquema de funcionamiento del RAID-5.....	230
Figura 8-23. Esquema de funcionamiento del RAID-0+1.....	230
Figura 8-24. Estructura del archivo .bmp.	239
Figura 8-25. Estructura de la cabecera .bmp.....	239
Figura 0-1. Dither usado para aliviar los efectos del error de cuantificación.....	253
Figura 0-2. El valor de amplitud es redondeado al escalón de cuantificación más cercano	254

ÍNDICE DE TABLAS

Tabla 1-1. Distintos márgenes de sensibilidad.	10
Tabla 2-1. Diámetros de altavoz más adecuados dependiendo de los diferentes diámetros de bobinas existentes.	46
Tabla 2-2. Tamaños recomendados del baffle de medida (norma IEC).	50
Tabla 4-1. Conexión del HDMI.	100
Tabla 4-2. Señal transportada en cada pin del DVI.	102
Tabla 4-3. Tabla comparativa entre las tecnologías más desarrolladas de monitores y televisión.	114
Tabla 5-1. Niveles posibles según el número de bits usados.	137
Tabla 5-2. Múltiplos posibles de frecuencia de muestreo en vídeo.	144
Tabla 5-3. Formatos de muestreo en televisión digital.	144
Tabla 6-1. Reducción de la tasa binaria con frecuencia de muestreo de 48 y 44,1 kHz.	161
Tabla 6-2. Comparativa de propiedades de bancos de filtros.	170
Tabla 7-1. Parámetros de los distintos estándares de televisión.	184
Tabla 7-2. Diferentes valores para la variable <i>chroma_format</i>	188
Tabla 8-1. Capacidades de almacenamiento y típicas duraciones de diferentes tipos de discos BD.	218
Tabla 8-2. Códec soportados por el Blu-ray.	219
Tabla 8-3. Especificaciones para los códecs de vídeo soportados por el sistema de Blu-ray.	220
Tabla 8-4. Cabecera BMP.	239
Tablas 0-1. Códigos VLC de luma y croma.	257
Tabla 0-2. Tipo de tabla a usar en función de las variables <i>intra_vic_format</i> y <i>macroblock_intra</i>	258
Tabla 0-3. Tabla B.14	260
Tabla 0-4. Ejemplo de un código escape	263

INTRODUCCIÓN

Los sistemas audiovisuales abarcan un amplio espectro de aplicación: desde la televisión, pasando por la radio, sistemas de telecomunicación, o incluso en el área del entretenimiento, salud o educación. Es por ello, por la importancia de estos sistemas, por lo que en este documento se pretende crear una base sólida de conocimiento sobre la materia pero a la vez para acercar al lector a la investigación sobre toda la ciencia que ésta envuelve.

Durante las últimas décadas la tecnología ha experimentado un notable avance y con ello los sistemas de sonido y vídeo e imagen han desarrollado características que los han hecho flexibles y robustos. Gracias a este avance, los sistemas audiovisuales permiten una interacción entre tecnología y persona. En este sentido cabe destacar la importancia de dichos sistemas en áreas como la domótica o del control de dispositivos mediante el reconocimiento sonoro o visual.

El desarrollo artístico también va ligado al avance tecnológico. Es verdad que la música por ejemplo no necesita de dispositivos de sonido para poderse producir, pero no podría llegar tan rápido ni tan lejos como lo hace actualmente. De hecho el avance en sí de los sistemas de sonido o vídeo ha influido directamente en la forma de tocar de los músicos o en peculiaridades de ciertos géneros musicales. Por ejemplo, el contrabajo sin amplificación servía únicamente de apoyo rítmico en el jazz, pero gracias a los sistemas electroacústicos su sonido podía ser amplificado, lo que daba a los intérpretes más posibilidades melódicas pudiendo transmitir mucho más a su audiencia. Otro ejemplo para ilustrar la interacción arte-tecnología es el Cine. Sin el avance de las cámaras no existiría siquiera esta forma de arte.

Hay por tanto una necesidad de avance e interacción entre diversos campos, ya sean artísticos, tecnológicos o humanos, en los que el audio y vídeo, juegan un papel clave de unión.

Vídeo y sonido siempre van de la mano. El porqué de esto se debe a la naturaleza propia del sentido de la vista y el oído. Bien es cierto que se puede ver sin oír y viceversa, pero para una completa captación de lo que nos rodea es necesario usar ambos sentidos. Con ellos somos capaces de comprender y analizar el entorno. Y esto mismo se traduce a los sistemas de audio y vídeo, el desarrollo de una tecnología se complementa con la otra llegando a compartir algoritmos, estándares de codificación o dispositivos electrónicos.

Los sistemas audiovisuales han ayudado a la expansión de la información. La globalización del mercado y el desarrollo de los sistemas de comunicación han permitido que la intercomunicación sea más rápida, sencilla y eficaz. En este punto se puede decir que, gracias al desarrollo tecnológico no sólo se establece interacción entre tecnología y persona, sino también entre diferentes tecnologías así como entre personas.

Actualmente el mundo del sonido y la imagen goza de un gran soporte. Agrupaciones como AES¹ o EBU² hacen encarecidos estudios sobre materia de televisión, radio, transmisión o sonido. Lo que hace de la materia una ciencia en un alto grado de desarrollo y avance.

El proyecto está enfocado para los alumnos de Grado en Sonido e Imagen, concretamente a aquellos que cursan la asignatura *Sistemas Audiovisuales*. Se ha tenido en cuenta la formación del alumno, es decir, qué asignaturas ha cursado antes que la que nos incumbe. Desde este punto de vista, para la fluida lectura de este documento se entiende que el alumno ha cursado las de Cálculo matemático, Electromagnetismo y Ondas, Señales y Sistemas, Fundamentos de Sonido e Imagen, Procesado Digital de la Señal, Propagación de Ondas y,

¹ Audio Engineering Society

² European Broadcasting Union

aunque no necesaria para la comprensión del documento, sí lo es para la asignatura en sí, Teoría de la Comunicación.

La asignatura *Sistemas Audiovisuales* está dividida en tres bloques temáticos. Este documento sirve como apoyo para los dos primeros bloques, el tercero no ha sido incluido dado que se puede encontrar suficientes recursos para su estudio en la página de la asignatura de Moodle.

Los bloques temáticos correspondientes a los temas:

1-Dispositivos de captación y reproducción de sonido e imagen.

2-Señales y formatos de audio y vídeo.

En cada bloque se sigue el orden espacial y temporal que obedecen las propias señales de sonido y vídeo. Esto es, el Tema 1 comienza con Micrófonos (1)³ para pasar a Altavoces (2) que es básicamente el recorrido principal de una onda sonora, el sistema lo ‘oye’ y después lo ‘dice’. En cuanto al vídeo, se ha seguido el mismo esquema, en primer lugar se ‘ve’ con la Cámara (3) para pasar a mostrar lo ‘visto’ con Proyector y Monitores (4).

El segundo bloque temático se centra en la forma que se le da a las señales y dónde y cómo se las almacena. Dado que toda señal en la realidad es ‘analógica’, es decir es continua, se parte desde ese punto. Sin embargo actualmente los sistemas analógicos están obsoletos, luego es necesaria una explicación del proceso para la Digitalización de audio y vídeo (5). Llegados a este punto se inicia al lector en la codificación de señales, que equivaldría al proceso de comunicarse en la vida real a través de un idioma determinado. Después el desarrollo del texto termina con los Soportes y Formatos de Almacenamiento (8), que no es más la forma de almacenar toda la información de audio y vídeo para usarla cuando sea necesario.

Se puede asemejar cada capítulo a una forma de comunicarse, recordar, ver u oír del ser humano. De hecho los sistemas audiovisuales están hechos a semejanza de los sentidos de la vista y el oído. Estándares de codificación como MPEG-1 o MPEG-2 se hacen siguiendo las características de percepción del oído humano.

Con los cuestionarios se sigue la misma línea estructural de cada capítulo, es decir que para responder una pregunta se necesita haber estudiado el apartado correspondiente a la pregunta anterior.

Uno de los principales problemas en la redacción del proyecto ha sido el simplificar una materia tan extensa y con infinitas aplicaciones como ya se ha explicado anteriormente. A pesar de que el documento es extenso, ha de saber el lector que en la corrección de cada capítulo se ha prescindido de multitud de información para sintetizar lo máximo posible y crear un texto fluido y sencillo. Con ayuda del material ya publicado en Moodle, se ha adaptado todo lo posible este documento a los requerimientos docentes de la asignatura.

Hay que remarcar que se ha cumplido con uno de los objetivos personales del autor: revisar y asentar los conocimientos adquiridos durante la carrera. Es necesario para elaborar un documento sólido dominar los conceptos de los que se habla, para lo que se necesita un estudio previo de todo lo que se va a publicar. Es por ello que en cada capítulo se han repasado toda la acústica, matemática, física y electrónica de la materia en cuestión. Además con ello se ha llegado a adquirir nuevos conocimientos, y esto es precisamente lo que se pretende con este proyecto, que el lector no sólo logre dominar los temas que aquí se explican sino que con algo de indagación se lleguen a comprender conceptos mucho más complejos para aplicarlos al mundo real.

Se han utilizado multitud de fuentes para hacer un texto abierto. Esto le permite adaptarse o complementar fácilmente la lectura de libros de mayor o menor dificultad.

³ Cada número entre paréntesis hace referencia al capítulo al que pertenece el texto.

TEMA 1

Dispositivos de captación y reproducción de Sonido e Imagen

1 MICRÓFONOS

1.1 Introducción a la señal de audio

Lo que percibimos como sonido es **energía acústica** que consiste en la fluctuación de ondas de presión en un medio físico, normalmente aire. Un ciclo completo de una onda de presión acústica consiste en medio ciclo de compresión de las moléculas del aire, seguido de medio ciclo de enrarecimiento. La velocidad de la fluctuación de la presión del aire se denomina **frecuencia de sonido**. El margen en frecuencia percibido por el sistema auditivo generalmente va de **20 Hz a 20 kHz**, para tonos puros.

Se llama **señal de audio** a la representación de los sonidos como señales eléctricas y se representa como una tensión o corriente que varía con el tiempo $V(t)$ tal y como se muestra en la Figura 1-1. Esto es así porque el voltaje o la corriente, fluctúan exactamente a la misma velocidad que la energía acústica que representan, con amplitudes proporcionales.

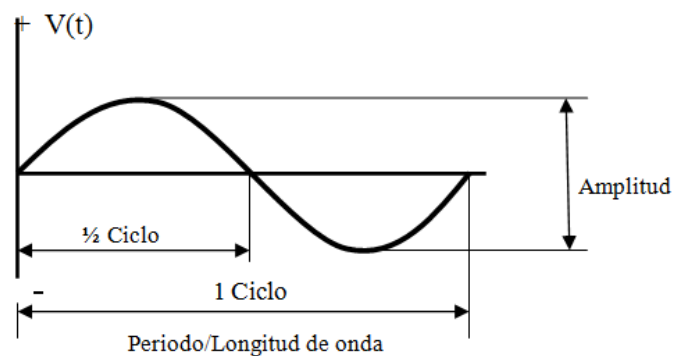


Figura 1-1. Representación de la señal de audio.

Qué es un sistema de sonido: Un sistema de sonido es una asociación de componentes electrónicos diseñado para amplificar la señal de audio tal y como se muestra en la Figura 1-2.

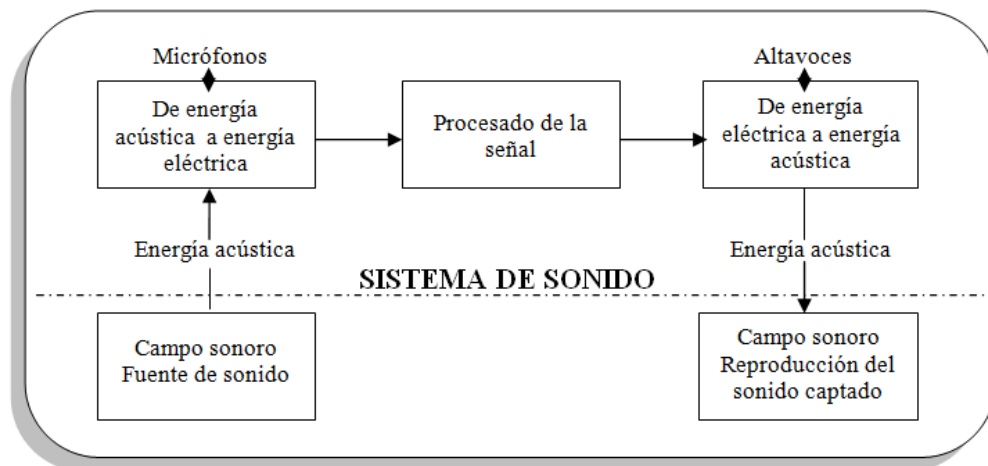


Figura 1-2. Esquema de un sistema de sonido.

Respuesta en frecuencia

La respuesta en frecuencia de un dispositivo describe la relación entre su entrada y su salida de acuerdo a la frecuencia de la señal y su amplitud. En otras palabras, nos muestra las frecuencias que el dispositivo deja pasar desde su entrada a su salida.

La respuesta en frecuencia depende de muchos factores: material de la membrana del micrófono, cableado, componentes como condensadores, resistencias, etc. Tanto micrófonos como en altavoces se diseñan para cubrir el máximo margen de frecuencias posibles.

Nivel de presión sonora

El decibelio describe la relación entre dos cantidades (en este ejemplo P1 y P2), generalmente potencias. Es usado para facilitar la compresión de magnitudes que sin esta relación serían tediosas de manejar. El decibelio deriva del Belio (B) que es la décima parte del decibelio (dB). La ecuación (1-1) calcula dicha relación:

$$10 \log (P1/P2) \text{ dB. (1-1)}$$

El término de nivel de presión sonora se refiere a la potencia sonora. No hay que confundir potencia sonora con nivel de presión sonora. Potencia sonora es la energía sonora total radiada por un altavoz u otra fuente sonora en todas las direcciones, mientras que la presión sonora es el nivel medido por unidad de área en una localización particular relativa a la fuente sonora. Al utilizar nivel de presión sonora la ecuación usada es dB_{SPL} . Esto es así porque la relación no está hecha para potencias sino para presiones. La ecuación (1-2) calcula el nivel de presión sonora:

$$SPL(dB_{SPL}) = 20 \log \frac{p_1}{p_{ref}} \text{ dB. (1-2)}$$

$$[p]=N/m^2$$

$$p_{ref} = 20 \mu Pa$$

Se usa 20 porque la potencia es proporcional a la presión al cuadrado. La presión de referencia es 20 μPa . El margen de nivel de presión sonora percibido por el sistema auditivo puede llegar a ser muy amplio, desde 0 dB_{NPS} hasta 120 dB_{NPS} .

Rango dinámico

Diferencia entre el nivel más alto y el más bajo de SPL que admite un dispositivo es el rango dinámico. Generalmente el nivel más bajo suele ser enmascarado por el ruido ambiente, por lo que en este caso el rango dinámico pasa a ser la diferencia de nivel entre al SPL mayor y el ruido de fondo.

El término de rango dinámico es también aplicado a los sistemas de sonido. Todo dispositivo tiene inherente ruido debido a los componentes que lo conforman. En este campo el rango dinámico es la diferencia entre el nivel pico de salida del sistema y el ruido electroacústico.

Por ejemplo, en un concierto de rock, un micrófono nos da un SPL de 40 dB debido a la audiencia, el viento, el ajetreo del escenario, etc., y 130 dB cuando la banda está tocando al máximo, con el cantante gritando. El rango dinámico en este caso calcularía con la expresión (1-3):

$$RD = (\text{Nivel pico}) - (\text{Nivel de ruido}) = 130 \text{ dB}_{SPL} - 40 \text{ dB}_{SPL} = 90 \text{ dB. (1-3)}$$

Los sistemas de sonido obviamente no funcionan a máxima potencia todo el tiempo (a esos 130 dB_{SPL}). Para saber los niveles a los que puede trabajar un dispositivo hay que fijarse en el nivel medio electrónico de línea usado para conciertos que es generalmente de 110 dB_{SPL} (1,23 V). El Headroom es la diferencia entre el nivel de pico del sistema y el nivel medio o nominal del mismo y se calcula con la ecuación (1-4):

$$HR = (\text{Nivel de pico}) - (\text{Nivel nominal}). (1-4)$$

El *Headroom* es una especificación que nos dice la capacidad de un sistema de sonido de aguantar picos de señal. Dados dos sistemas que operen al mismo nivel nominal, el que mejor HR tenga será el que mejor soporte los picos en la señal que está pasando a través de él.

1.2 Definición y características de los micrófonos

Un micrófono es un dispositivo que transforma la energía acústica (sonido) en energía eléctrica (señal de audio).

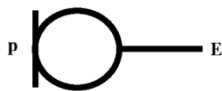


Figura 1-3. Símbolo del micrófono. 'p'-presión incidente, 'E'-señal eléctrica

Es lo que llamaremos un transductor acústico-eléctrico. Está formado por una membrana o diafragma que recibe la onda de presión acústica y la envía a un convertidor acústico-mecánico (**TAM**) y éste a su vez traslada el movimiento a un transductor mecánico-eléctrico (**TME**) del que obtenemos la señal de audio. Un micrófono se puede dividir por tanto en dos transductores tal y como se muestra en la Figura 1-4.

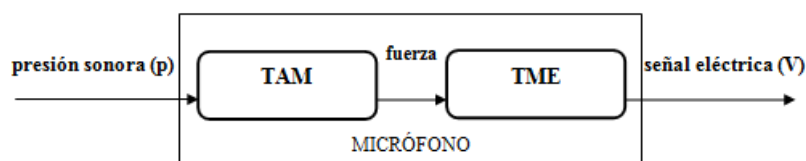


Figura 1-4. Transductores que conforman el micrófono.

El **TAM** convierte las variaciones de presión acústica (p) de las ondas sonoras en vibraciones de un elemento móvil llamado diafragma del micrófono. La respuesta del TAM al medio determina la respuesta directiva y la respuesta en frecuencia del micrófono.

El **TME** convierte las vibraciones del diafragma en variaciones de tensión (V) y corriente, es el llamado motor del micrófono. El TME determina la sensibilidad del micrófono ante el nivel de presión sonora y la adaptación eléctrica como etapa de salida que se conecta a las siguientes etapas del sistema de sonido.

Sensibilidad

Definimos la sensibilidad como la presión sonora mínima que hay que ejercer sobre el diafragma para que proporcione una señal eléctrica. Esta característica nos indica la capacidad del micrófono para captar sonidos débiles o de poca intensidad. Los micrófonos generan pequeños niveles de señal (señal de micro) si lo comparamos con otros dispositivos, como mesas de mezclas (señal de línea). Por ello los micrófonos requieren una **preamplificación** de su señal de salida. El nivel de salida del micrófono siempre es especificado con respecto a un nivel de presión sonora a su entrada, generalmente a 1000 Hz.

La **sensibilidad** de un micrófono (S), que se calcula con la ecuación (1-5) es el módulo de la relación entre la tensión eléctrica proporcionada a la salida del micrófono en circuito abierto ($V_{c.a.}$) y la presión acústica incidente en el micrófono (p).

$$S = \left| \frac{V_{c.a.}}{p} \right| \quad (1-5)$$

La medida está realizada para 1 kHz, en campo libre, en el eje del micrófono y a 1m de distancia de la fuente sonora.

Las unidades de la sensibilidad son mV/Pa ó mV/ μ bar. Hay que recordar que 1 Pa = 10 μ bar. El TEM es el que determina la sensibilidad del micrófono ante el nivel de presión sonora. Pero la sensibilidad siempre suele venir indicada en dB.

$$S = 20 \log \frac{S[V/Pa]}{1 V/Pa} < 0 \quad \text{cuyas unidades son dB referenciadas a } 1 V/Pa$$

$$S = 20 \log \frac{S[V/\mu bar]}{1 V/\mu bar} < 0 \quad \text{cuyas unidades son dB referenciadas a } 1 V/\mu bar$$

La sensibilidad en dB es menor que 0 porque los micrófonos siempre entregan muy poco voltaje. Los micrófonos de uso profesional suelen tener una sensibilidad entre 1 mV/Pa y 10 mV/Pa. En la Tabla 1-1 se expresa estos márgenes de sensibilidad en las diferentes unidades.

Tabla 1-1. Distintos márgenes de sensibilidad.

S [mV/Pa]	S [mV/μbar]	S [dBre 1 V/Pa]	S [dBre 1 V/μbar]
1	0.1	-60	-80
10	1	-40	-60

Distorsión

La distorsión es el comportamiento no lineal del transductor, es decir, la deformación de la forma de onda que ha captado a su entrada. Este comportamiento se puede dar tanto en la respuesta en frecuencia del micrófono como en su respuesta temporal.

La distorsión puede ser debida tanto a causas internas del diseño electroacústico del dispositivo, como a causas externas como la sobrecarga de presión o saturación del micrófono debido a niveles altos de presión sonora. Se cree generalmente que la sobrecarga en un sistema de sonido está causada por el micrófono. Y en realidad es muy difícil que un micrófono sobrecargue. Profesionalmente los micrófonos son capaces de resistir niveles de presión sonora de hasta más 140 dB_{SPL} sin sobrecargarse. Para los distintos tipos de micrófonos, que se verán más adelante, una mala alimentación puede hacer que el SPL soportado por el micrófono disminuya. Los micrófonos pueden tener un preamplificador eléctrico interno que también puede ser saturado, por eso algunos micrófonos suelen llevar un atenuador eléctrico de señal conmutable, llamado PAD, antes del preamplificador.

La **respuesta transitoria** también influye en la distorsión microfónica. La respuesta transitoria de un micrófono es la capacidad de reproducir fielmente los cambios bruscos de presión acústica, producidos generalmente por fuentes sonoras como baterías, pianos y todo sonido causado por un suceso sonoro instantáneo. En señal hablada este factor no tiene demasiada importancia ya que los cambios de presión son fácilmente captados por los micrófonos.

Respuesta en frecuencia

La respuesta en frecuencia de un micrófono, $S(f)$, muestra la sensibilidad que tiene a diferentes frecuencias. Se puede decir que, una respuesta en frecuencia ideal sería aquella en la que el micrófono es capaz de trasladar el nivel de presión sonora al mismo nivel de señal de audio sin importar la frecuencia, es decir, la respuesta ideal es plana de 20 Hz a 20 kHz. Con la ecuación (1-6) se calcula la respuesta en cada frecuencia:

$$S(f) = 20 \log \frac{S(f)}{S(1 \text{ kHz})} \text{ dB. (1-6)}$$

Se expresa en dB referidos al valor de sensibilidad medida en campo libre a 1 kHz por ello en todas las gráficas de la respuesta en frecuencia pasan por 0 dB a dicha frecuencia, como se muestra en la Figura 1-5.

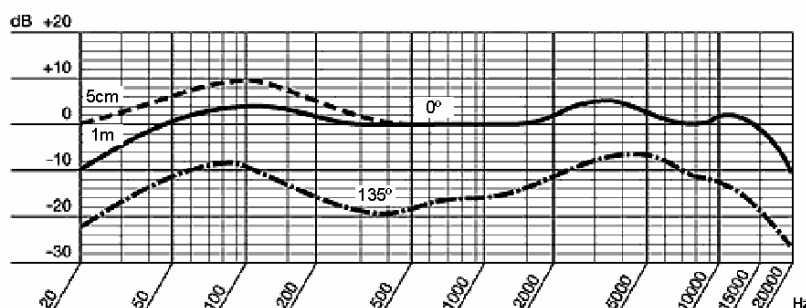


Figura 1-5. Respuesta en frecuencia de un micrófono. Se especifica la respuesta axial para $x = 1\text{m}$ y para $x = 5\text{cm}$. También la respuesta para un ángulo de incidencia de la señal de 135°

En micrófonos para aplicaciones vocales, suele aparecer la respuesta en frecuencia en campo libre, para diferentes distancias y ángulos, entre la fuente y el micrófono. Pero no hay ningún micrófono con respuesta plana ideal, de hecho, en ocasiones la respuesta no plana del mismo puede incurrir a una mejora a frecuencias muy útiles para el habla o la música, mejorando su inteligibilidad o parámetros subjetivos relativos al cantante o tipo de música. De todas maneras hay que ser capaz de suplir estas carencias del los micrófonos para no hacer variar la tonalidad u otras características que definen un determinado suceso auditivo.

Un efecto producido por determinados micrófonos es denominado **efecto proximidad** (ver Figura 1-5), mediante el cual las bajas frecuencias sufren un incremento de nivel que puede superar los 16 dB. Esto se produce cuando el micrófono se sitúa a menos de 60 cm (campo próximo) de la fuente sonora. Esto puede causar sobrecarga en el preamplificador efecto totalmente negativo si lo que se desea es aumentar la inteligibilidad de la señal hablada.

Directividad

Es la capacidad que tienen los micrófonos para entregar una tensión eléctrica diferente para cada ángulo de incidencia asociado a la onda de presión de entrada. En la práctica la mayoría de micrófonos son directivos, es decir, entregan más tensión si la onda de presión les llega de frente que lateralmente. Son omnidireccionales los micrófonos que dan la misma tensión de salida independientemente del ángulo de llegada del sonido. La directividad puede beneficiar la captación de sonido ya que puede ayudar en el rechazo del ruido de fondo o de la reverberación, reduce la posibilidad de que se produzca realimentación acústica (el típico pitido momentáneo que se oye en los conciertos) y permite independizar la captación de diferentes fuentes simultáneamente. Para determinar la directividad de los micrófonos se tiene en cuenta el ángulo de elevación, θ . En la Figura 1-6 se muestra las coordenadas usadas entre la fuente y el micrófono para la obtención de un **diagrama polar de directividad**, el cual nos muestra la respuesta del micrófono según el ángulo de llegada de la onda de presión sonora.

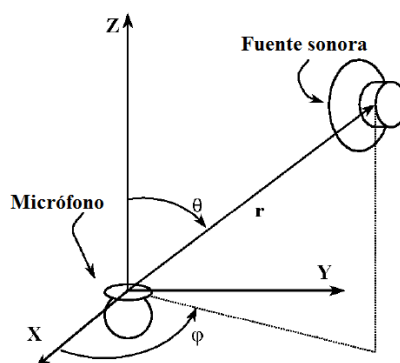


Figura 1-6. Coordenadas esféricas para la determinación de la directividad de un micrófono

Donde la directividad se calcula con la ecuación (1-7):

$$D(\theta, \varphi) = \frac{S(\theta, \varphi)}{S(0,0)} \quad (1-7)$$

El **diagrama polar de directividad** se mide en cámara anicónica (campo libre, sin reflexiones) utilizando como excitación bandas de ruido rosa en tercios de octava desde 125 Hz. En la Figura 1-7 se muestra un ejemplo de un diagrama polar de directividad y en la Figura 1-8 se muestra un esquema de las condiciones de medida de un diagrama polar de directividad.

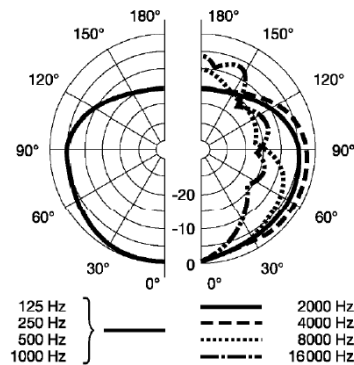


Figura 1-7. Diagrama polar de Directividad a diferentes frecuencias

La directividad acaba siendo una relación de tensiones que entrega el micrófono que por relación entre las ecuaciones (1-5) y (1-7) se obtiene la relación entre tensiones de la ecuación (1-8):

$$D(\theta) = \frac{S(\theta)}{S(0)} = \frac{E(\theta)/p(r)}{E(0)/p(r)} = \frac{E(\theta)}{E(0)} \quad (1-8)$$

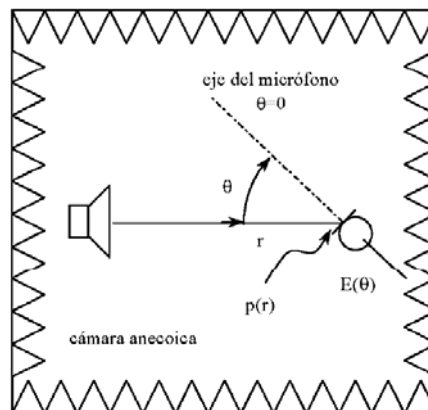


Figura 1-8. Condiciones en las que se realiza la medida del diagrama polar de Directividad.

Según la directividad se puede hacer la siguiente clasificación de micrófonos:

-Omnidireccional: Nunca son usados en refuerzo sonoro ya que no proporcionan protección ante la realimentación acústica. En este tipo de micrófonos se genera el mismo nivel de señal eléctrica, independientemente de la dirección de la que provenga el sonido. Esta característica es más fácil que se cumpla a bajas frecuencias, ya que a frecuencias altas predomina la captación por la parte frontal.

Son útiles en los siguientes casos:

- Captación del sonido en todas las direcciones.
- Captación de reverberaciones en locales, cámaras etc.
- Exclusión máxima del ruido mecánico generado por viento etc.
- Respuesta amplia en las frecuencias más bajas, sobre todo con micrófonos capacitivos.

-Directivos: Todos los micrófonos directivos, en mayor o en menor medida sirven para:

- Rechazar al máximo la acústica que tenga el recinto donde se realiza la toma.
- Rechazar el ruido de fondo.
- Utilizar técnicas especiales de grabación con parejas de micrófonos (estéreo coincidente).

- Captación de sonidos lejanos.

Hay diferentes tipos de micrófonos directivos:

-Bidireccional: tienen sensibilidad máxima para los sonidos que inciden frontalmente sobre el diafragma, ya sea por la cara anterior o por la posterior.

-Cardioide: Sólo recogen sonidos frontalmente. Como se puede observar en la figura, la sensibilidad es máxima para los sonidos de proveniencia frontal y va disminuyendo a medida que varía el ángulo de incidencia sobre el diafragma.

-Hipercardioide: Capta mejor las señales que inciden frontalmente, pero también capta señales por la parte posterior.

En la Figura 1-9 se muestran los diagramas de los tipos de micrófonos según su directividad.

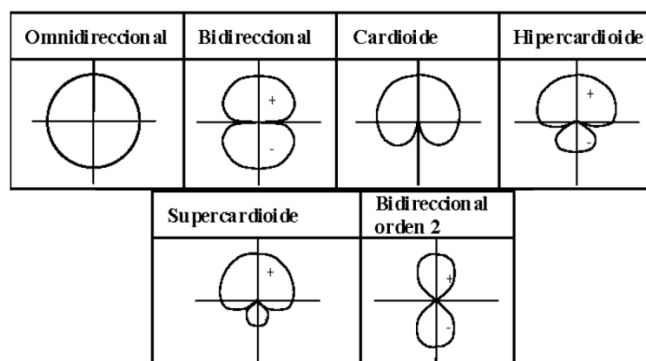


Figura 1-9. Patrones más comunes de directividad.

-Ancho de haz: o ángulo de captación, es el ángulo abarcado en el diagrama polar de directividad para una cierta pérdida de sensibilidad. Las pérdidas más comunes suelen ser las de -3 dB y -6dB como se muestra en la Figura 1-10. Sus anchos de haz son respectivamente **130°** y **180°**.

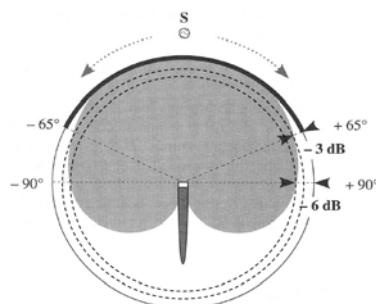


Figura 1-10. Ancho de haz para -3 dB y -6 dB.

-Parámetros de directividad: sirven para analizar la directividad de manera global y no ángulo por ángulo.

i) **Factor de directividad axial, Q_{AX} .** El factor de directividad queda definido por la ecuación (1-9):

$$Q(\theta, \varphi) = \frac{S^2(\theta, \varphi)}{\langle S^2(\theta, \varphi) \rangle} \quad (1-9)$$

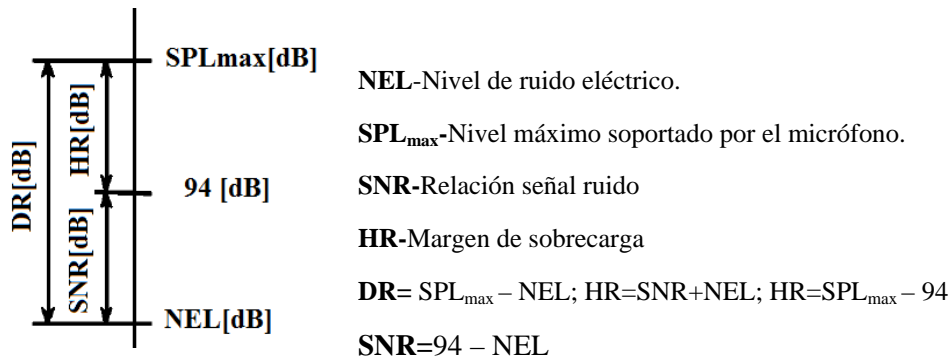
Siendo $S^2(\theta, \varphi)$ la sensibilidad al cuadrado (energía captada por el micrófono en campo difuso) y $\langle S^2(\theta, \varphi) \rangle$ el promedio de la sensibilidad al cuadrado (promedio en todo el espacio de la energía captada por su equivalente omnidireccional, también en campo difuso).

El factor de directividad axial queda definido como $Q_{AX} = Q(0, 0)$. Si $Q_{AX}=1$ el micrófono es omnidireccional pero si $Q_{AX}>1$ el micrófono es directivo.

ii) **Eficiencia de energía aleatoria, $REE=1/Q_{AX}$** : si $REE=1$ el micrófono es omnidireccional, si $REE<1$ el micrófono es directivo. El logaritmo de este factor es el **Índice de directividad**= $10\log(REE)$

Margen dinámico, DR

Como ya se ha comentado con anterioridad, los micrófonos tienen un extraordinario margen dinámico, con valores comprendidos entre 90 dB y 120 dB. Esto es porque admiten mucha presión acústica sin distorsionar y entregan un nivel de ruido muy bajo. Pero a pesar de su gran margen dinámico entregan tensiones muy bajas del orden de unos pocos milivoltios, pudiendo ser enmascarada por el ruido captado por los cables a los que se conecta o al de los preamplificadores.



Impedancia

La impedancia nominal, Z_0 , es la impedancia de salida del generador equivalente de Thevenin que representa el micrófono. Debemos tener en cuenta que la impedancia de salida del micrófono puede ser, como máximo, de un tercio de la del equipo al que se conecta. Un valor típico de la impedancia de un micrófono está comprendido entre 200 y 600 ohmios. En este rango de valores se mueven los micrófonos de baja impedancia, generalmente los más usados en refuerzo sonoro y grabación. Los micrófonos de alta impedancia suelen rondar resistencias de más de 25 kΩ.

Algunos modelos de micrófonos ya incorporan un conmutador para seleccionar la impedancia. Si no fuese así, podemos utilizar adaptadores de impedancia.

La **impedancia de carga mínima (Z_i)** es la que garantiza que el micrófono transfiera toda la tensión posible sin pérdidas. Se recomienda que la impedancia de carga mínima sea 3 veces la impedancia nominal del micrófono.

En baja impedancia se pueden utilizar cables largos, mientras que en alta no es conveniente debido a las pérdidas que estos generan.

1.3 Tipos de micrófonos

Los micrófonos se pueden dividir en diferentes tipos dependiendo de su transductor acústico-mecánico (TAM) o su transductor mecánico-eléctrico (TME).

1.3.1 Según su TAM

El TAM determina como recibe el micrófono las ondas de presión y como es la conversión a señales de vibración mecánica. Para ser más precisos, relaciona la presión acústica externa al micrófono con la fuerza, f_p , que impulsa su diafragma. Este es el transductor que determina la directividad del micrófono.

a) Micrófonos de presión

En este tipo de micrófonos la vibración del diafragma es proporcional a la presión acústica externa al micrófono que se enfrenta a la cara externa del mismo. Esto es así porque la cara interna del diafragma está aislada del

exterior por la cavidad que constituye la capsula microfónica. Esta configuración hace que la cara externa tenga en un instante determinado de tiempo una presión $P_0 + p(t)$, siendo P_0 la presión atmosférica y $p(t)$ la presión acústica. La vibración del diafragma está marcada por la fuerza f_D que lo impulsa, y que a su vez está originada por la diferencia de presión entre ambas caras de diafragma. Dicha fuerza se calcula como muestra la ecuación (1-10):

$$f_D(t) = [p_1 - p_2] \cdot S_D = p(t) \cdot S_D \text{ [Pa} \cdot \text{m}^2\text{]}. \quad (1-10)$$

Siendo S_D la superficie efectiva del diafragma. En la Figura 1-11 se puede ver a grandes rasgos las características del micrófono de presión.

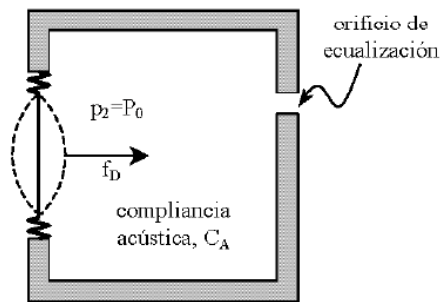


Figura 1-11. TAM del micrófono de presión

El aislamiento de la capsula microfónica impide la movilidad del diafragma en ausencia ondas acústicas, pero a su vez puede provocar su ruptura. Esto se puede producir en caso de que haya grandes cambios barométricos, a los que el diafragma respondería moviéndose produciéndose la ruptura del mismo. Por ello todos los micrófonos de presión incorporan un orificio de ecualización o igualación de presiones que comunica la cavidad interna con el exterior de forma que las presiones interna y externa del micrófono se pueden igualar. El tubo de ecualización cumple la misma función en el micrófono que la trompa de Eustaquio en nuestro oído.

Directividad: Los micrófonos de presión son omnidireccionales. Dado que solo es activa la cara externa del diafragma si el sonido incide desde atrás por difracción rodea la capsula microfónica y acaba empujando al diafragma desde la cara externa ya que no puede acceder a la cara interna. Esto solo ocurrirá cuando las ondas traseras rodeen sin dificultad a la capsula, lo cual sucederá siempre que el diámetro del diafragma sea mucho menor que la longitud de onda de la señal acústica incidente ($D \ll \lambda$). Luego a mayor frecuencia la omnidireccionalidad disminuirá, por ello los micrófonos de presión de gran tamaño no funcionan bien en altas frecuencias. Para cubrir todo el margen de audio se admiten tamaños de diafragma no mayores a media pulgada. En la Figura 1-12 se puede apreciar el porqué de la directividad omnidireccional.

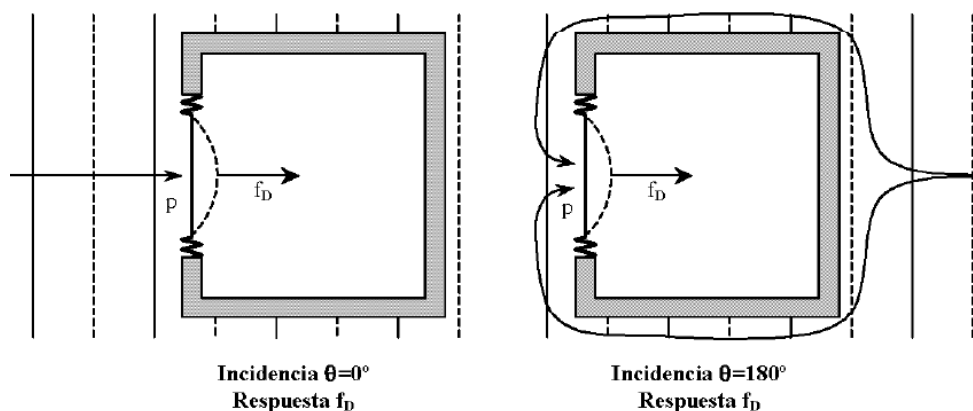


Figura 1-12. Cuando la onda incide con $\theta=180^\circ$ rodea la capsula y empuja al diafragma de igual forma que si incidiese desde $\theta=0^\circ$.

b) Micrófonos de gradiente de presión

En este tipo de micrófonos las ondas acústicas pueden acceder por las dos caras del diafragma y su respuesta es proporcional a la diferencia de presiones que existe entre ambas caras del diafragma. Esta diferencia de presiones es de carácter diferencial en baja frecuencia, de ahí el término 'gradiente'. Es decir, la fuerza mecánica f_D que impulsa al diafragma, es proporcional a la diferencia de presiones existente entre ambas caras del diafragma. Dependiendo del ángulo, θ , de llegada de la onda sonora, los caminos que ésta recorre hacia cada cara del diafragma serán más o menos parecidos. Para los ángulos 90° y 270° la diferencia de presión entre las dos caras del diafragma es nula luego no habrá movimiento del mismo. Según esto se puede deducir que la directividad de este tipo de micrófonos corresponde al patrón bidireccional. En la Figura 1-13 se muestra la estructura física del micrófono de gradiente de presión.

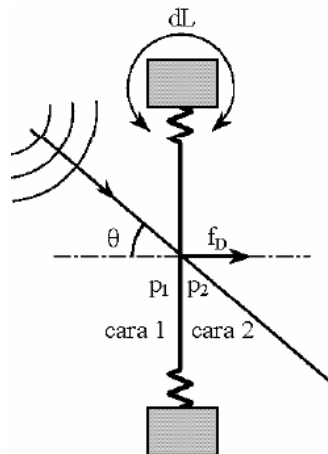


Figura 1-13. TAM del micrófono de gradiente.

Es en este tipo de micrófonos en los que se produce el efecto de proximidad de que se habló anteriormente. Éste se da concretamente cuando el micrófono está sometido al campo próximo de ondas de presión esféricas y a baja frecuencia. Esto es así porque el gradiente de presión en las ondas esféricas es mayor que en las ondas planas. En la Figura 1-14 se detalla una respuesta en frecuencia en la que se refleja el efecto proximidad.

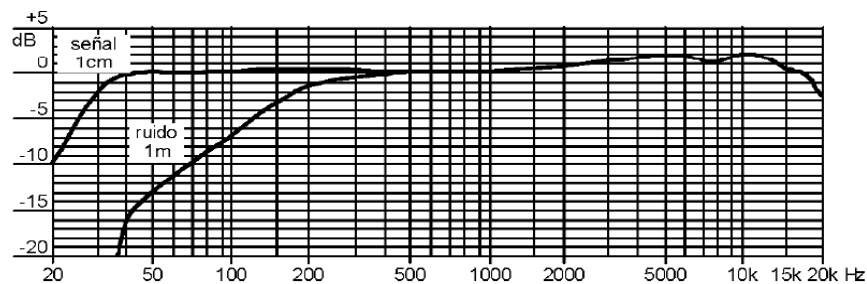


Figura 1-14. Efecto proximidad de los micrófonos de gradiente.

Directividad: En la práctica, los micrófonos de gradiente de presión se diseñan para que tengan directividad en forma de ocho y consecuentemente se trata de evitar la zona de alta frecuencia. Para ello se diseñan con un diafragma bastante pequeño, para que el primer mínimo de captación se sitúe en altas frecuencias, fuera de la zona de trabajo del transductor, alrededor de 8kHz para micrófonos vocales.

Que el diagrama tenga esta forma de 8, quiere decir que la tensión entregada en ambos casos será de signo distinto, luego hay que tener cuidado a la hora de mezclar micrófonos de gradiente cuando estos capten señal de la misma fuente, ya que si no se alinean los dos lóbulos positivos se puede producir una cancelación en la respuesta global. En la Figura 1-15 se muestran los diagramas polares de los micrófonos de gradiente en bajas y altas frecuencias.

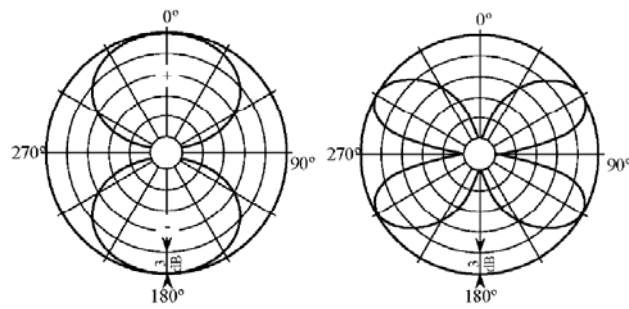


Figura 1-15. Diagrama polar de directividad en baja frecuencia (izqda.) y en alta frecuencia (dcha.)

c) Micrófonos combinados de presión y gradiente

Se denominan así a aquellos receptores electoracústicos que responden mecánicamente a la presión acústica externa al micrófono y además al gradiente de presión existente entre ambas caras del diafragma. Su respuesta es equivalente a una suma de las salidas eléctricas de un micrófono de presión puro y de un micrófono de gradiente de presión puro. Los micrófonos combinados tienen una cavidad interna que se comunica con el exterior por medio de unas aberturas de tamaño relativamente pequeño para ofrecer cierta resistencia al paso de la onda de presión acústica. Pertenecen a la familia cardioide. En la Figura 1-16 se muestra la estructura de los micrófonos combinados de gradiente y presión.

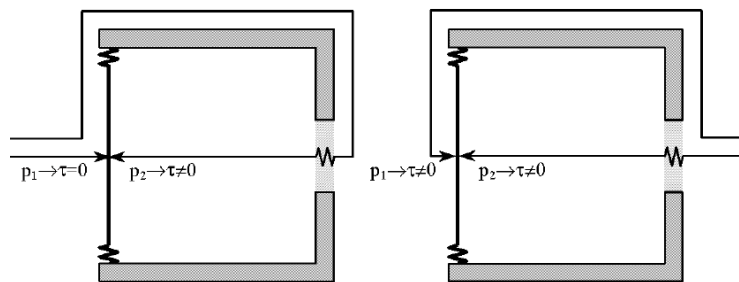


Figura 1-16. TMA del micrófono combinado de gradiente y presión.

Cuando la onda de presión llega de frente, Figura 1-16 izquierda, hay parte de ella que incide en la cara posterior del diafragma y parte se deriva a la parte posterior del mismo tras atravesar las aberturas y atenuarse. En este caso la diferencia de presión entre ambas caras del diafragma será grande, luego fuerza de impulso del diafragma f_D será grande.

Cuando la onda de presión llega desde atrás, Figura 1-16 derecha, la diferencia de presiones entre la cara anterior y posterior del diafragma será menor, luego f_D también lo será.

Directividad: Pertenecen a la familia cardioide, manteniéndose con una directividad estable para bajas frecuencias. En agudos tiende a estropearse el diagrama polar. A medida que la onda es más esférica la configuración tomada por estos micrófonos es la bidireccional.

1.3.2 Según su TME

a) Micrófonos dinámicos de bobina móvil

El micrófono dinámico es, posiblemente, el más ampliamente usado para grabación de sonido, difusión y sistemas de dirección pública. Es muy robusto y puede ser diseñado para proporcionar altas prestaciones. A diferencia del micrófono capacitivo, que veremos después, no requiere de una fuente de alimentación y su impedancia de salida es baja, de modo que no necesita etapas buffer para su acoplamiento a un cable.

Se basa en el principio de inducción electromagnética (son la versión dual de los altavoces de bobina móvil) según el cual si un hilo conductor se mueve dentro de un campo magnético, en el conductor se inducirá un voltaje de acuerdo con la ecuación (1-11):

$$E = BLv \text{ [V]}. (1-11)$$

E = potencial inducido; **B** = densidad de flujo magnético; **L** = longitud del conductor; **v** = velocidad del movimiento.

El funcionamiento es el siguiente: la onda acústica incide en el diafragma provocando su vibración afectando directamente a la bobina a la que está unido. La bobina se mueve dentro de un campo magnético, lo que induce una corriente eléctrica proporcional a la onda de presión acústica incidente. Es resistente a diversos factores ambientales lo que lo hace muy útil para captación de sonido en exteriores. En la Figura 1-17 se muestra la estructura de un micrófono de bobina móvil.

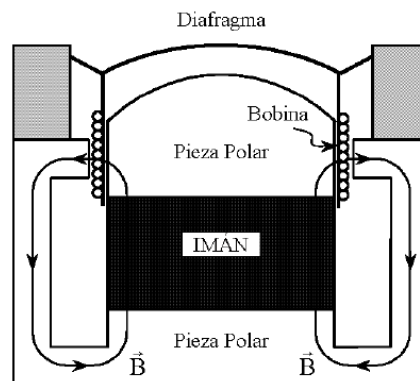


Figura 1-17. Micrófono dinámico de bobina móvil funcionando como micrófono de presión.

Su respuesta en frecuencia es irregular. Pueden tener diferentes directividades, dependiendo de cuál sea su modo de operación acústico mecánico. Pueden ser omnidireccionales o cardioides. Su impedancia nominal es resistiva y suele tener valores en torno a los 100 Ω .

Construcción: los micrófonos de bobina móvil están constituidos por una cápsula con carcasa de plástico o goma, donde se aloja el transductor y que se encuentra aislada mecánicamente de la estructura externa, para evitar que se transmita el ruido de manipulación. El diafragma suele ser de plástico, el circuito magnético de hierro y el imán cerámico. Pueden llevar filtros eléctricos que aminoran el efecto proximidad. En la Figura 1-18 se muestran los elementos estructurales del micrófono de bobina móvil.

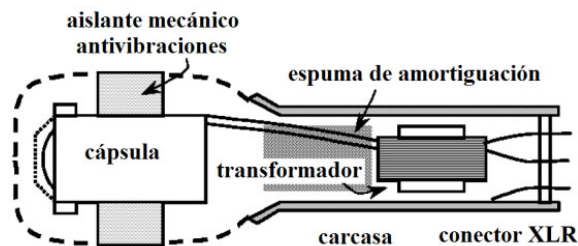


Figura 1-18. Esquema de un micrófono dinámico de bobina móvil.

b) Micrófonos dinámicos de cinta

Su funcionamiento es similar al de los micrófonos de gradiente. Este TME consiste en una cinta corrugada sujeta en sus extremos está sometida a un campo magnético fijo. Cuando el sonido incide en la cinta ésta vibra y el campo magnético induce una tensión en ella. El voltaje inducido es muy pequeño y la impedancia muy baja, de 0,1 Ω , luego los micrófonos de cinta incorporan un transformador que amplifica dicha tensión y aísla su pequeña impedancia.

Los micrófonos de cinta son muy frágiles, la cinta se puede estropear con facilidad por un soplo, luego no es muy utilizado. Su respuesta en frecuencia es bastante irregular con mala respuesta en agudos y bastante marcado el efecto proximidad. Generalmente presentan una directividad bidireccional aunque pueden adoptar diversas directividades de la familia cardioide, pero nunca pueden tomar la configuración omnidireccional. En la Figura 1-19 se muestra la estructura funcional del micrófono dinámico.

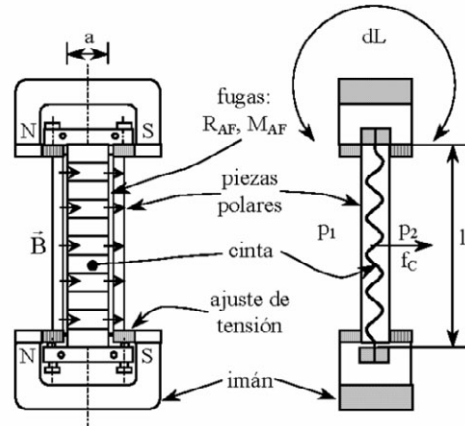


Figura 1-19. Micrófono dinámico de cinta funcionando como micrófono de gradiente de presión.

Se usan normalmente como micrófonos vocales o para sistemas de pares estéreo; también en retransmisiones de tipo deportivo, pero no se suelen usar en exteriores, y en caso de que este fuese su fin, tendrían que llevar protección antiviento.

Construcción: la cinta suele ser de unos pocos centímetros y su anchura de unos pocos milímetros, con un espesor de micrómetros y generalmente es de aluminio. Están separadas unas pocas micras de las piezas polares. Tiene unas corrugaciones para que su comportamiento sea lo más lineal posible. El imán y las piezas polares son grandes por lo que son micrófonos pesados.

c) Micrófonos electrostáticos de condensador

Después del micrófono de bobina móvil, es el más utilizado. Funciona por el principio del condensador variable. El diafragma es el elemento móvil y constituye una de las placas del condensador. La otra placa es fija. Cuando incide sobre el micrófono la onda de presión, el diafragma se mueve variando la capacidad del condensador y por tanto también la carga interna, produciéndose una corriente de audio sobre el circuito eléctrico de carga, que normalmente es un preamplificador de gran impedancia de entrada.

En Figura 1-20 T_r es la tensión radial del diafragma y E_{ca} es la tensión en circuito abierto que entrega el micrófono.

Constan de una sensibilidad muy alta y margen dinámico elevado de hasta 150 dB o más. Su respuesta en frecuencia es extremadamente plana, dependiendo de su respuesta en agudos. Normalmente micrófonos con diafragmas de media pulgada llegan a cubrir sin problemas hasta los 20 kHz.

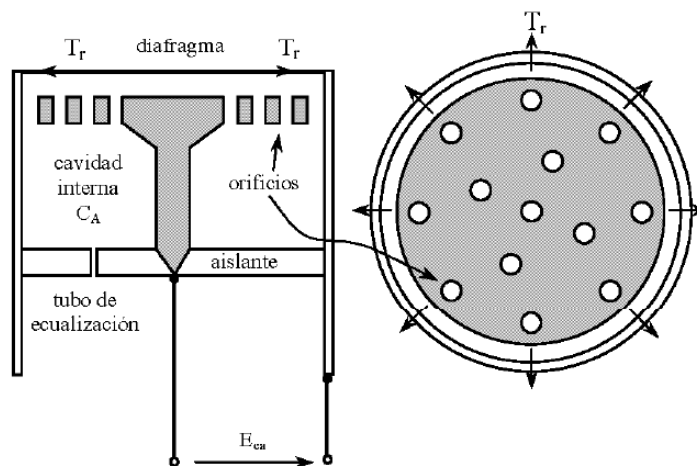


Figura 1-20. Micrófono electrostático de condensador funcionando como micrófono de presión.

Es necesario conectarlos a preamplificadores de impedancias de entrada muy altas, del orden de los $G\Omega$ para reproducir sin problemas en audiofrecuencia. Normalmente se fabrican como omnidireccionales o cardioides pero no como bidireccionales. A diferencia de los dinámicos, mantienen la directividad con la frecuencia en casi todo el margen de trabajo.

Su impedancia nominal tiene valores de aproximadamente 100 ohmios.

Construcción: no se sigue exactamente el mismo proceso dependiendo si el micrófono es de medida, vocal o instrumental. La carcasa y la armadura fija perforada se realizan de aleaciones metálicas (Cu, Fe, Ni, Al) y el diafragma tensado suele ser de acero, níquel o aluminio, con espesores de $e=10\mu m$ y una separación entre placas de $d_0=10$ micras. La capacidad de los condensadores suele rondar los 10 picofaradios.

En caso de que se trate de un micrófono de uso vocal, la placa fija se suele recubrir de un plástico aislante para evitar que la condensación o los golpes pongan en contacto eléctrico ambos electrodos y el micrófono se rompa.

La tensión de polarización de la cápsula, E_p , oscila entre los 50V y los 200V. En micrófonos de medida esta tensión se proporciona por una fuente continua conectada a un circuito eléctrico independiente; en micrófonos de uso vocal esa alimentación se mezcla con la señal de audio. Es lo que se denomina la alimentación '*Phantom*' o fantasma. Sus valores están normalizados entre 12V-48V. En la Figura 1-21 se muestra el circuito electrónico de la alimentación *Phantom*.

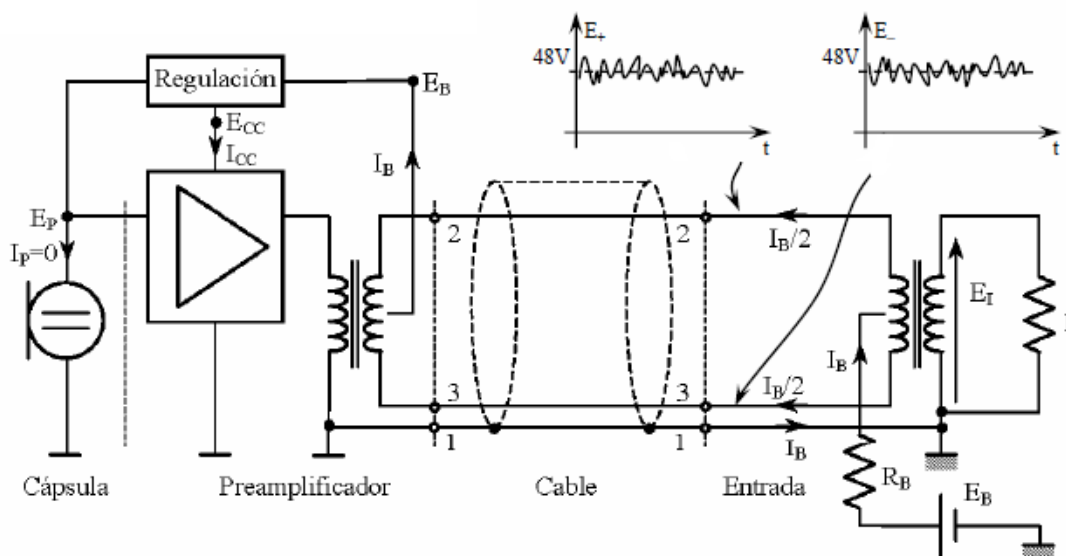


Figura 1-21. Alimentación *Phantom* para micrófonos de condensador.

d) Micrófonos electrostáticos de electret o prepolarizados

Los micrófonos de electret son equivalentes a los micrófonos electrostáticos de condensador, con la diferencia de que no necesitan una tensión continua de polarización E_p , ya que el condensador lleva una carga interna fija, de ahí la denominación de prepolarizados.

Un material electret es aquel con polarización permanente. Éstos suelen ser de tipo plástico como el Teflón y el Mylar.

Su sensibilidad suele ser alta, con un amplio margen dinámico. Aun siendo muy pequeños pueden conservar una alta sensibilidad lo que permite su miniaturización. Su respuesta en frecuencia es similar a la de los electrostáticos de condensador. En cuanto a la directividad, puede tomar cualquier tipo menos la configuración bidireccional. Su impedancia nominal es la del preamplificador al que va unido. En la Figura 1-22 se muestra de forma esquemática la estructura de los dos tipos de micrófonos, electret o polarizados.

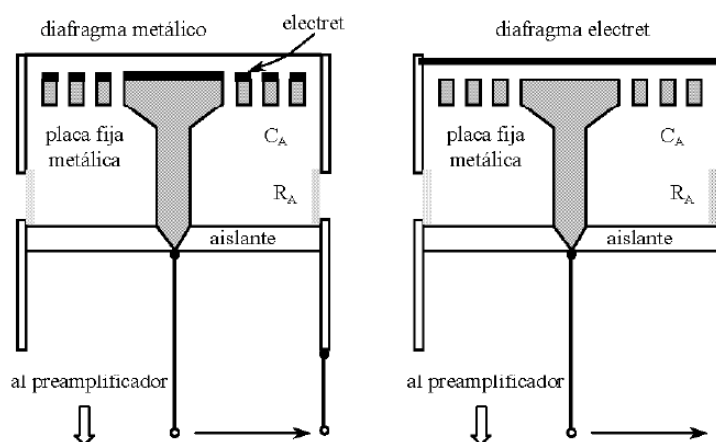


Figura 1-22. Esquema constructivo de un micrófono de electret con diafragma metálico (izqda.) y con diafragma electret (dcha.)

Construcción: está conformado por la cápsula microfónica y el preamplificador. El diafragma puede ser metálico o de material electret, aunque estos últimos tienen una peor respuesta.

La protección para que las placas no se toquen es mayor que en los micrófonos de condensador, ya que si esto ocurriese el micrófono no se rompería. La impedancia de salida es baja, y el preamplificador se encarga de la adaptación de impedancias. Aunque no sea necesaria tensión para la polarización si es necesaria para la alimentación del preamplificador, unos 10V, aunque los hay que funcionan con pequeñas pilas de 1,5V.

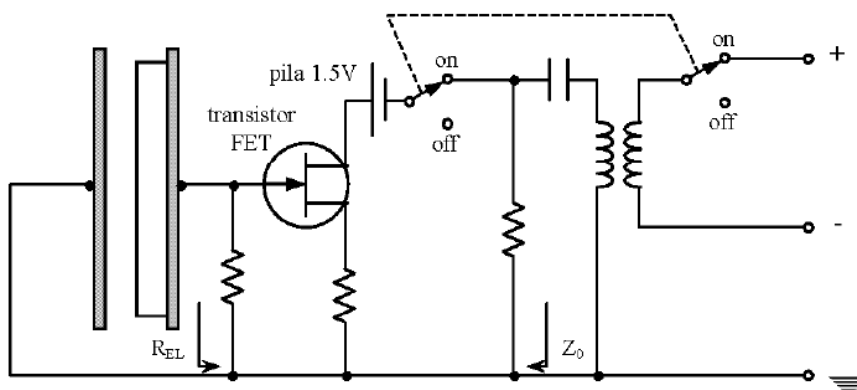


Figura 1-23. Esquema de un micrófono electret alimentado con una pila de 1,5V.

Suelen ser usados como micrófonos para cámaras de video, equipos domésticos de audio o de multimedia informática.

e) Micrófonos de carbón

Se basan en las propiedades de ciertos materiales que varían su resistencia al someterlos a deformaciones. Constan de una membrana que cierra una cavidad en la que se sitúan granos de carbón. En cada una de las caras se sitúan unos terminales eléctricos, y el circuito se cierra con una tensión constante.

En reposo, obtenemos una corriente constante, pero en cuanto una onda sonora incide sobre el diafragma, los granos de carbón se comprimen, varía su resistencia y obtenemos una corriente variable en función de la presión de la fuente sonora.

Tienen una calidad baja y su coste también resulta bajo. Se utilizan en aplicaciones como porteros automáticos o en telefonía.

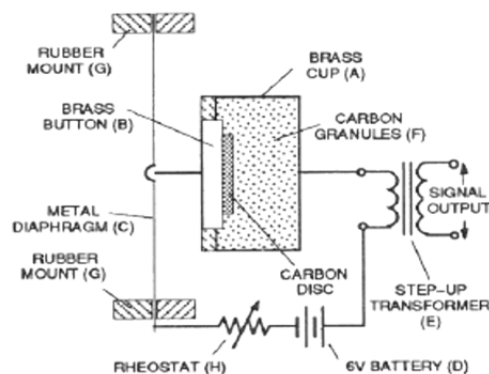


Figura 1-24. Construcción de un micrófono de carbón.

f) Micrófonos de piezoeléctricos

Se basan en la propiedad que tienen algunos materiales cerámicos o cristalinos de generar una tensión eléctrica cuando se someten a una compresión. De esta forma, las vibraciones que se producen en el diafragma se transmiten a un material de este tipo que, al deformarse, produce una tensión eléctrica que es la de salida del micrófono.

Suelen tener una impedancia de salida muy alta (1 MΩ), por lo que, normalmente, necesitan adaptación para su conexión a cualquier equipo. Además, las propiedades de los materiales piezoeléctricos se alteran con la temperatura, lo que resulta una característica poco deseable.

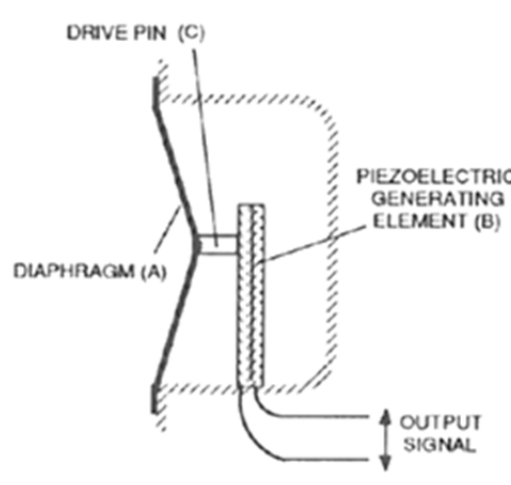


Figura 1-25. Construcción de un micrófono piezoeléctrico.

g) Otros Micrófonos

-Micrófonos de doble diafragma: constituido por dos diafragmas y una placa fija que actúan como dos electrodos metálicos de dos condensadores variables. El usuario puede elegir la directividad del micrófono mediante conmutadores. Los diafragmas no son totalmente independientes ya que la placa fija está perforada y lo que afecta a uno afecta, de forma similar, al otro. Son utilizados con exclusividad en grabaciones de estudio. Poseen elevada sensibilidad y margen dinámico, así como una respuesta en frecuencia muy plana.

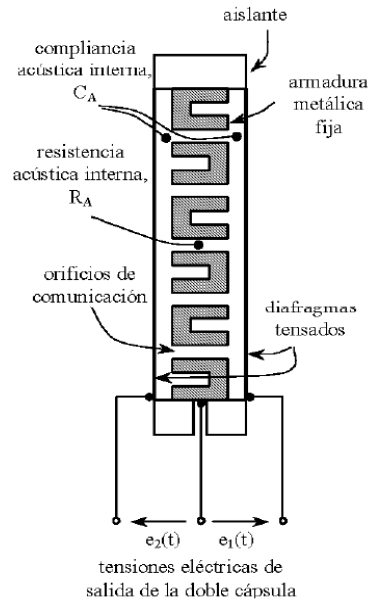


Figura 1-26. Esquema constructivo de un micrófono de doble diafragma.

-Micrófonos superdirectivos: de cañón o con paraboloide: cañón está constituido por una capsula directiva (cardioide o hipercardioide) a la que se antepone un tubo de interferencia. El tubo de interferencia es un cilindro con aperturas laterales para permitir el paso de la onda acústica.



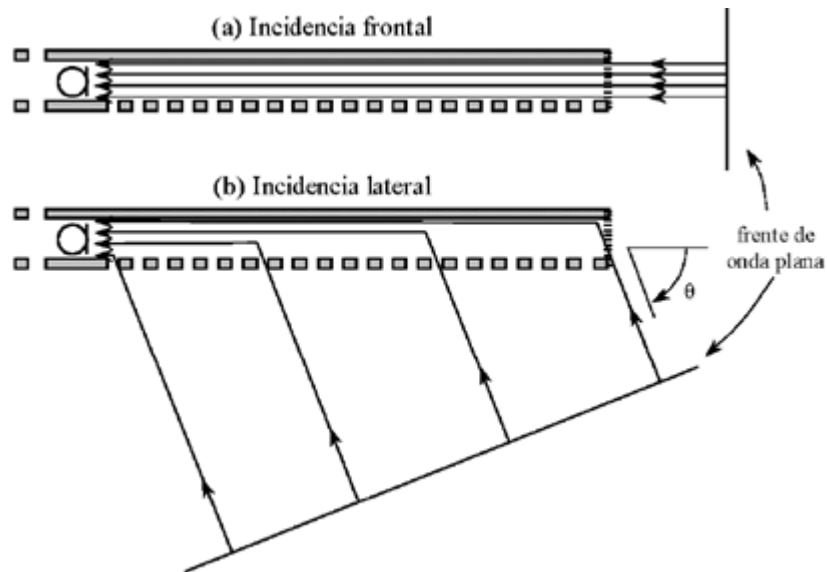


Figura 1-27. Comportamiento de un micrófono de cañón.

Su curva de directividad presenta un elevado lóbulo lateral y trasero que se acentúa con la frecuencia. Son usados para 'acercar' una fuente distante, lo cual lo hace muy útil para eventos deportivos o cualquier espectáculo con fuentes lejanas, como el cine.

Los micrófonos con paraboloides están compuestos por una cápsula microfónica y una pantalla reflectante en forma de paraboloide. La cápsula microfónica puede ser directiva y se debe situar en el foco del paraboloide tal y como se muestra en la Figura 1-28.

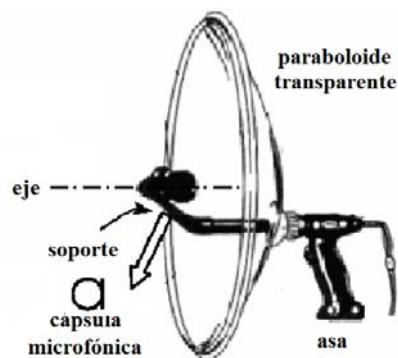


Figura 1-28. Micrófono con paraboloide

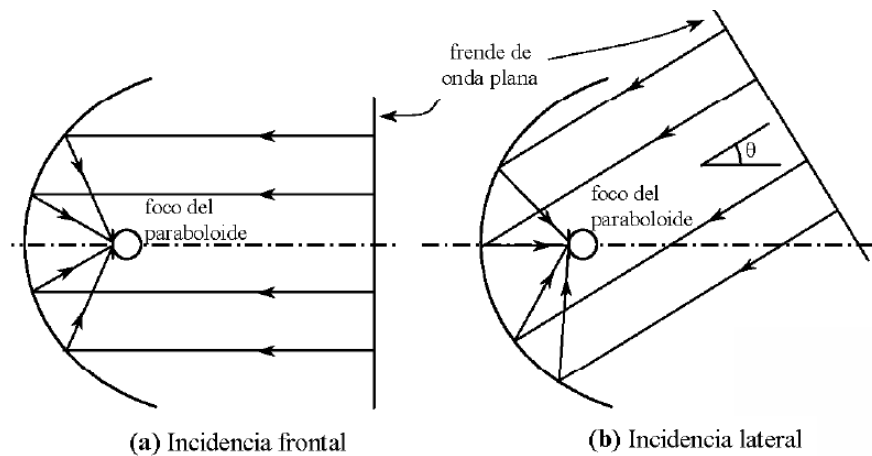


Figura 1-29. Captación del sonido de la cápsula microfónica de los micrófonos con paraboloide

Estos micrófonos consiguen una amplificación de las ondas acústicas en su eje y además una directividad muy elevada.

Micrófonos lavalier, de corbata o solapa: se colocan en la corbata o solapa de un locutor o colgado de un cordón. Están diseñados especialmente para la captación de voz.

Además evitan la captación del ruido producido por el roce de la ropa (no tienen efecto proximidad). Son de muy pequeño tamaño, por lo que son micrófonos de tipo electret.

Micrófonos inalámbricos: son, a pequeña escala, un sistema de retransmisión FM. Son muy utilizados en televisión. Su esquema funcional se muestra en la Figura 1-30.

La microfonía inalámbrica tiene el problema del desvanecimiento de la señal, denominado *fading* o *drop-out*. Esto puede deberse a las condiciones del canal multitrayecto emisor-receptor o por algún obstáculo. Para evitar esto se incorporan sistemas de diversidad espacial, que consiste en que el receptor incorpora dos o más antenas, y mediante un conmutador, controlado por el nivel de portadora recibido, se toma la señal válida. Esto se muestra de forma esquemática en la Figura 1-31.

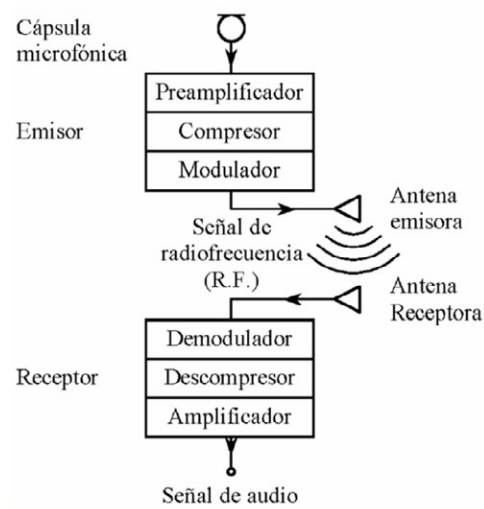


Figura 1-30. Esquema de funcionamiento de los micrófonos inalámbricos.

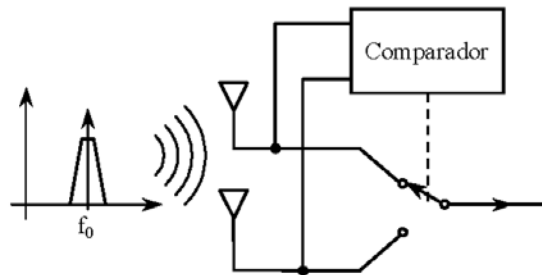


Figura 1-31. Diversidad espacial activa o 'verdadera'.

Los diferentes tipos de antenas usadas en microfonía inalámbrica se muestran en la Figura 1-32 :

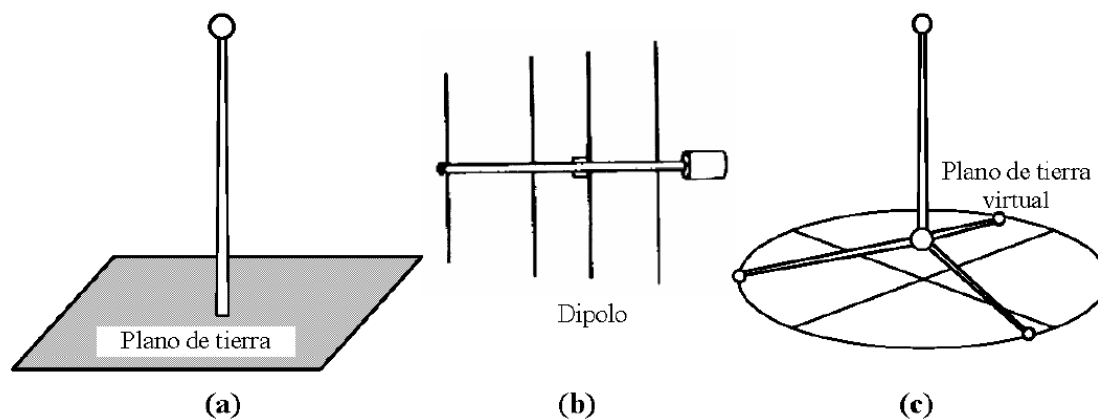


Figura 1-32. Antenas receptoras. a) Monopolo en $\lambda/4$ con plano de tierra. b) Dipolo en $\lambda/2$ sin plano de tierra. c) Monopolo en $\lambda/4$ con plano de tierra virtual.

1.4 Conexionado

Primero hay que diferenciar los tipos de señales que se manejan en audio. Se pueden diferenciar en:

i) **Señales de micro** o de bajo nivel, cuyo voltaje varía solo entre μV y mV . Son señales muy sensibles a contaminarse con interferencias y son las que tienen en su salida los micrófonos.

ii) **Señales de línea**, son las utilizadas en todos los equipos de audio como mesas de mezclas, ecualizadores, etc. Su rango de voltaje puede llegar hasta los 24,5 V.

Luego el principal reto cuando queremos conectar un micrófono a cualquier equipo, es protegerlo de las interferencias electromagnéticas que pueda captar el cable al que va unido y procurar que esos milivoltios de los que se disponen, puedan ser amplificados para reproducir con la mayor fiabilidad posible lo que se captó.

Lo más común a la salida de los micrófonos es encontrarse con un conector denominado **Cannon** o **XLR-3** macho. En la Figura 1-33 se muestra las partes del conector XLR al completo (macho-hembra).



Figura 1-33. Macho (13) y hembra (25) del XLR.

Carcasa externa metálica (*Rugged metal Shell*): Alberga, protege y apantalla los contactos. Además sirve de guía para la conexión macho-hembra.

Aislante (*insulator*): soporte y protección de los contactos metálicos, aislándolos uno del otro.

Contactos (*pins*): elementos metálicos que en la interconexión sirven de paso a la corriente eléctrica.

Guías (*guide*): muescas que sirven para que la introducción del conector se produzca en una orientación determinada.

Abrazadera (*guide pins*): asegura mediante compresión la sujeción del cable al conector, bien sea a la carcasa o al aislante.

El XLR es usado para señales de micro y de línea, es muy robusto y su conexión es muy sencilla. Tiene un periodo de vida muy alto.

1.4.1 Tipos de conexiones

Hay dos maneras de enviar la señal de audio:

i) Asimétrica: Es el envío de una variación de tensión con respecto a un potencial de referencia fija, es decir, variaciones de señal respecto a masa.

El cable usado comúnmente en el envío de señal asimétrica el **coaxial**. En la Figura 1-34 se representa la estructura del mismo.

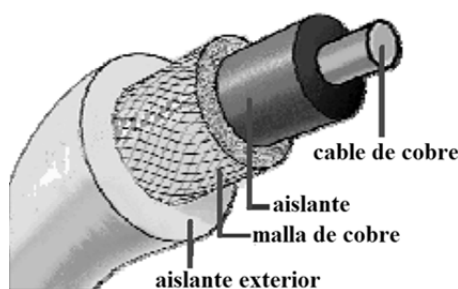


Figura 1-34. Estructura del cable coaxial.

El cable coaxial suele ser usado en instalaciones de consumo pero no en instalaciones profesionales de audio. La principal razón es la poca protección frente a interferencias electromagnéticas (EMI) que tiene. Las EMI provocan una corriente en la malla que se suma a la señal de audio y esto, en su reproducción, se traduce en ruido. Para intentar solventar el problema del ruido se han de usar cables cortos y trabajar en entornos poco ruidosos, además de tener una señal de audio de gran amplitud. Ninguna de estas condiciones se cumple en el audio profesional.

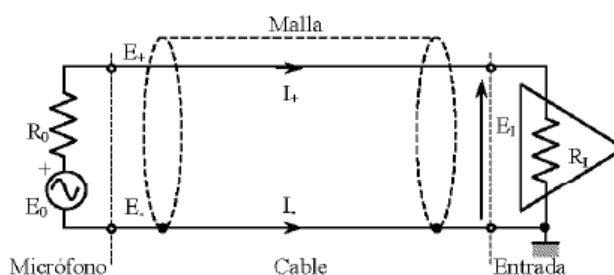


Figura 1-35. Esquema eléctrico de la conexión asimétrica.

ii) Conexiones balanceadas: Por la que se envía una señal diferencial, es decir, $V_A - V_B$, en la que V_A es una mitad de la señal de audio y V_B es la otra mitad pero invertida. De esta manera, el amplificador diferencial nos entregará la misma señal que teníamos a la entrada. En la Figura 1-36 se detalla el proceso y en la Figura 1-38 se muestra el proceso de envío de esta señal.

V_G : señal a enviar

$$V_s = V_A - V_B = (+V_G/2) - (-V_G/2)$$

$$V_s = V_G$$

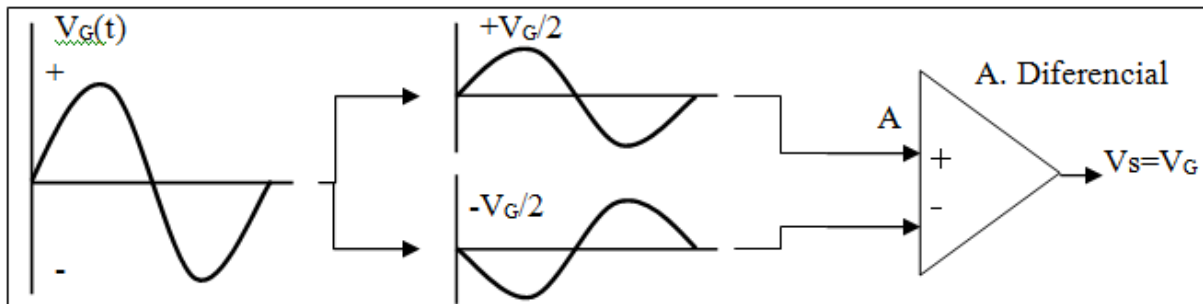


Figura 1-36. Esquema del envío balanceado de señal.

Para realizar este envío se utiliza un cable denominado **par trenzado apantallado** mostrado en la Figura 1-37.

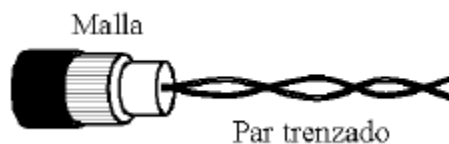


Figura 1-37. Par trenzado apantallado.

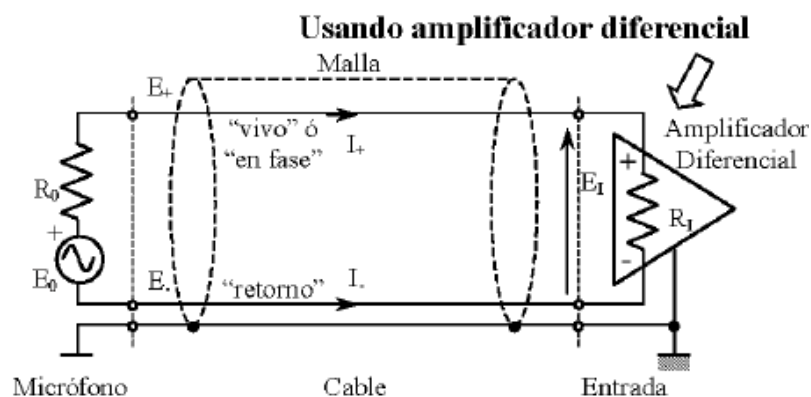


Figura 1-38. Esquema eléctrico del envío balanceado.

La principal ventaja de este tipo de cable es que, aparte de minimizar las EMI por la pantalla (malla de cobre), en caso de que la señal llegue a ser contaminada por ruido, afectará de la misma manera a V_A y V_B , por lo que el amplificador de señal, al realizar $V_A - V_B$ eliminará a su vez el ruido.

$$V_s = (+V_G/2 + \text{ruido}) - (-V_G/2 + \text{ruido}) = V_G$$

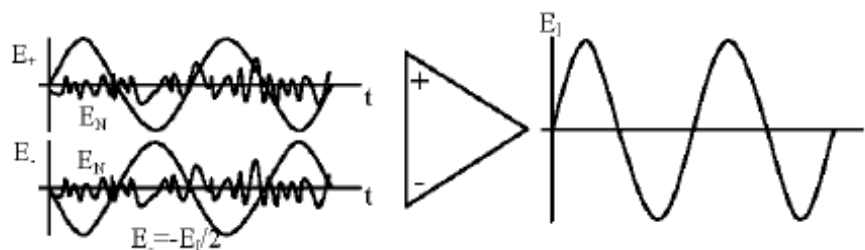


Figura 1-39. Rechazo del ruido en conexión balanceada.

La pantalla en el par trenzado no tiene relación alguna con la señal balanceada de audio que transmite, por ello la protección ante las EMI es mucho más efectiva. Es la conexión preferente en audio profesional, y es imprescindible usarla cuando hay largas tiradas de cable.

El conector XLR tiene además la ventaja de poder llevar tanto señal asimétrica como balanceada. A continuación se detalla según normativa qué señal han de recibir cada uno de los contactos del XLR:



- 1-Malla: tierra
- 2-Audio, en fase
- 3-Audio, retorno

1.5 Cuestionario

1-Un ciclo completo de una onda de presión acústica consiste en:

- a) Medio ciclo de compresión de las moléculas del aire, seguido de otro ciclo de compresión.
- b) Medio ciclo de compresión de las moléculas del aire, seguido de otro ciclo de refracción.
- c) Medio ciclo de refracción de las moléculas del aire, seguido de otro ciclo de refracción.
- d) Ninguna de las anteriores es correcta.

2-La analogía del mundo eléctrico con las ondas acústicas se justifica porque:

- a) El sonido, al igual que el voltaje y la corriente, consiste en una fluctuación de electrones con un periodo, frecuencia y longitud de onda que los definen.
- b) Es una manera de simplificar la acústica, siempre se comete un error.
- c) Voltaje o corriente fluctúan exactamente a la misma velocidad que la energía acústica que representan, con amplitudes proporcionales.
- d) Ninguna de las anteriores es correcta.

3-Si la referencia de nivel de potencia acústica es 1pW ¿Cuál es el nivel de potencia sonora de una fuente que emite a 1 mW ?

- a) 120 dB.
- b) 85 dB.
- c) -60 dB.
- d) Ninguna de las anteriores es correcta.

4-La sensibilidad de un micrófono es:

- a) El grado de distorsión que tienen.
- b) La capacidad de los micrófonos de no captar los sonidos inesperados.
- c) La capacidad que tiene para transformar la presión sonora en tensión eléctrica.
- d) Ninguna de las anteriores es cierta.

5-Un micrófono tiene una sensibilidad de 5 mV/Pa . Calcule la sensibilidad que tiene en dB re 1V/mbar .

6-Un micrófono tiene una sensibilidad de 5mV/Pa , calcule la tensión a su salida cuando se encuentra sometido a un nivel de presión sonora de $47\text{ dB}_{\text{SPL}}$ a la frecuencia de 1 kHz (suponer incidencia axial). Respuesta en mV.

7-Señale la frase correcta:

- a) El patrón de directividad bidireccional representa que el micrófono sólo capta el sonido que se produce en torno a su eje.
- b) El patrón de directividad bidireccional es característico solamente de los micrófonos de presión.
- c) El patrón de directividad bidireccional establece la misma polaridad eléctrica independientemente del ángulo de llegada del sonido.
- d) Ninguna de las anteriores es cierta.

8-La directividad es la capacidad que tienen los micrófonos para entregar una tensión eléctrica diferente para cada ángulo de incidencia asociado a la onda de presión de entra.

- a) Verdadero.

b) Falso.

9- Selecciona las afirmaciones correctas:

(Selecciona una o más de una)

- a) En decibelios, la directividad siempre será mayor que 1, ya que muestra la capacidad de ganancia del micrófono.
- b) El diagrama polar de directividad muestra la respuesta del micrófono según el ángulo de llegada de la onda de presión sonora para 1kHz.
- c) Los micrófonos omnidireccionales sirven para captar con mayor fiabilidad la onda de presión principal, aislándose del ruido de fondo y de la reverberación.
- d) Todas las afirmaciones anteriores son incorrectas.

10-Selecciona la afirmación correcta:

- a) En los micrófonos de presión la vibración del diafragma es proporcional a la presión acústica externa al micrófono que se enfrenta a la cara externa del mismo.
- b) La ecuación de la fuerza que mueve el diafragma es: $f_D(t)=[p_1-p_2] \cdot S_D=p(t) \cdot S_D$.
- c) En los micrófonos de presión la vibración de las ondas acústicas pueden acceder por las dos caras del diafragma, y su respuesta es proporcional a la diferencia de presiones que existe entre ambas caras del diafragma.
- d) La ecuación de la fuerza que mueve el diafragma es $f_D(t)=[p_1-p_2] \cdot S_D=p(t) \cdot S_D$ siendo p_0 la referencia de ondas sonoras.

11-El efecto proximidad:

- a) Se produce cuando la captación de la señal se hace en el campo próximo de la fuente sonora. Se da principalmente cuando se captan ondas esféricas
- b) En una característica propia de los micrófonos de electret o prepolarizados.
- c) Se produce cuando al preamplificador le llegan señales originadas por ondas de presión esféricas de alta frecuencia.
- d) Ninguna de las anteriores es cierta.

12-Los micrófonos de condensador:

- a) Son muy usados ya que no necesitan tensión continua de alimentación, lo que hace que su coste disminuya y su facilidad de uso sea mayor que la de otros micrófonos.
- b) Precisan de una tensión de alimentación para la polarización de la cápsula microfónica. La tensión de polarización viene proporcionada por un circuito denominado 'Phantom' y toma valores de 12V a 48V.
- c) Se basan en el principio de la inducción electromagnética: la parte móvil del condensador, que se mueve en un campo magnético fijo, que provoca una variación de corriente.
- d) Se basan en las propiedades de ciertos materiales que varían su resistencia al someterlos a deformaciones.

13- Una de las principales diferencias entre las señales de micro y de línea es:

- a) Que las señales de línea solo se utilizan para telefonía mientras que las de micro están reservadas para el campo de audio.
- b) La magnitud de su tensión.

- c) La magnitud de su sensibilidad.
- d) No tienen diferencia, son denominaciones distintas pero son la misma señal.

14-La señal asimétrica:

- a) Es la mejor protegida está ante interferencias y ruido.
- b) Es la usada en audio profesional.
- c) Es la usada en audio de consumo.
- d) Ninguna de las anteriores es correcta.

15- Los sistemas de microfonía inalámbricos son, a pequeña escala, sistemas de retransmisión FM.

- a) Verdadero.
- b) Falso.

16- Asociar cada pin del conector XLR a su función en la transmisión de señales balanceadas.

Opciones: Malla (tierra), audio en contra-fase (retorno) y audio en fase.

- Pin 1:
- Pin 2:
- Pin 3:

17- El par trenzado apantallado:

- a) Tiene una malla que favorece a la captación de las EMI.
- b) No es lo suficientemente robusto como para transmitir señales balanceadas.
- c) Usa un solo hilo conductor lo suficientemente robusto como para la transmisión de señal asimétrica y balanceada.
- d) Ninguna de las anteriores es correcta.

18- Los dos tipos de micrófonos más utilizados son:

- a) Dinámico y de condensador.
- b) Dinámico y PZM.
- c) De cinta y dinámico.

19-Un micrófono es un:

- a) Transductor de entrada.
- b) Transductor de salida.
- c) Ninguna de las anteriores.

20-Un micrófono:

- a) Muestrea la señal de audio.
- b) Convierte las variaciones de presión sonora en una señal de audio.
- c) Convierte de analógico a digital.
- d) Convierte la señal de audio en variaciones de presión sonora.

2 ALTAVOCES

Altavoz es un término general para describir una amplia variedad de transductores que convierten la energía eléctrica en energía acústica o sonido.

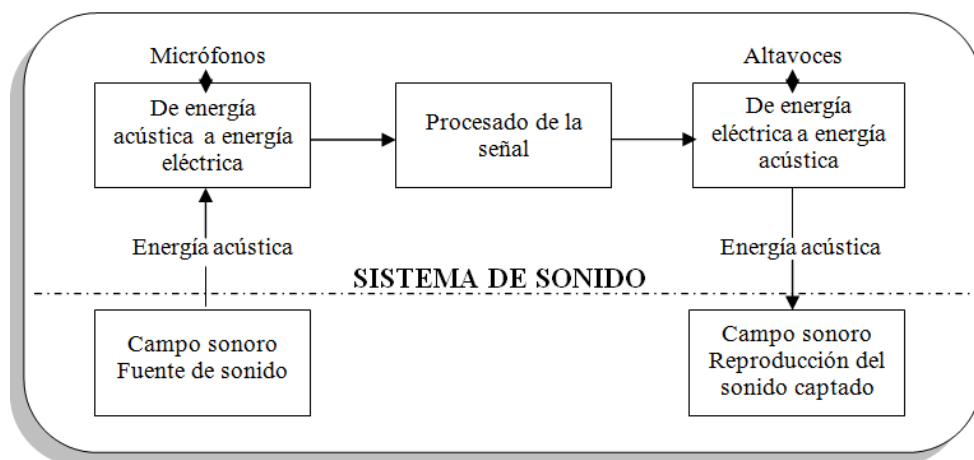


Figura 2-1. Esquema de un sistema de sonido.

En un sistema de sonido, los altavoces juegan un rol muy importante como la unión final entre la fuente de sonido y la audiencia. Además si nos fijamos en el esquema de un sistema de sonido, el paso previo es el procesado de la señal de audio. Este puede ser tanto digital como analógico. En la Figura 2-1 se muestra el esquema general de un sistema de sonido.

Antes de comentar las características de los altavoces, se comentará brevemente los dispositivos involucrados para que llegue la señal de audio a los mismos.

Mesas de mezclas

Una mesa de mezclas es un complejo sistema de audio, que no sólo preamplifica y asigna señales de entrada a diferentes salidas, también permite una variedad de diferentes mezclas, un enrutamiento y procesamiento especial de la señal de audio.

Las entradas a las mesas de mezclas son de micro o de línea. Todas las entradas van de un bus de mezcla al control general de niveles del que se obtiene la señal de salida de la consola. Se tiene un número limitado de entradas que se asocian a un bus determinado. La asignación se puede hacer mediante control de niveles o con ‘interruptores’ de señal (envío o no envío).

La consola permite ecualizar y filtrar la señal de cada canal de entrada. A las mesas de mezclas también se les pueden conectar compresores, procesadores de efectos, etc. En la Figura 2-2 se muestra un ejemplo de una mesa Yamaha.



Figura 2-2. Mesa de mezclas modelo M7CL de Yamaha.

Amplificadores de potencia

Como su propio nombre indica se encargan de amplificar la potencia de una señal de audio. Los amplificadores profesionales en apariencia son sencillos, sólo amplifican, no se ocupan de filtrar ni ninguna otra tarea. En ocasiones, las mesas de mezclas pueden incluir ya en su interior un amplificador de potencia.



Figura 2-3. Amplificador de potencia modelo D100 de DAS Audio.

Tienen que entregar altas potencias a un tipo de cargas muy concretas, altavoces de $8\ \Omega$ o $4\ \Omega$, o a sistemas determinados durante largos periodos de tiempo, luego han de ser muy robustos. Tienen una etapa de amplificación de tensión, en la que se convierte la entrada de línea a decenas de voltios; a la que sigue la etapa de corriente, en la que se genera la potencia requerida. En la Figura 2-4

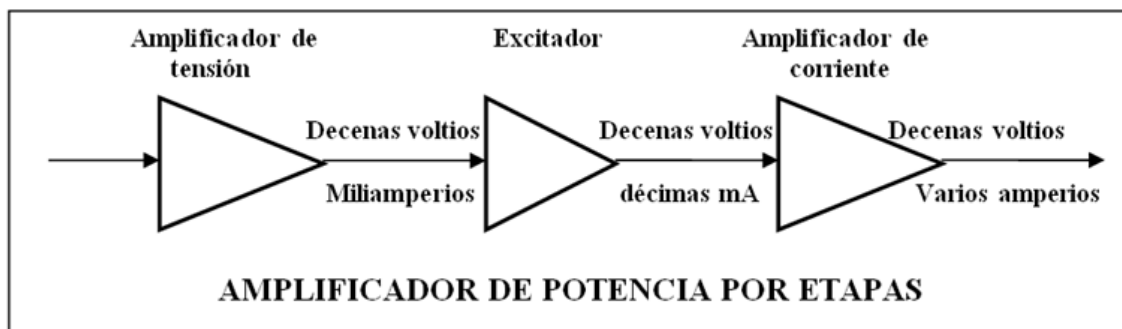


Figura 2-4. Etapas de un amplificador de potencia.

Etapa de filtrado

Debido a las características de los elementos que conforman el altavoz, es difícil encontrar sistemas que radien en toda la banda de audio con un solo altavoz. Por ello se diseñan **filtros de cruce** (*crossover*) en los sistemas de altavoces, para dividir el espectro en varias bandas y alimentar a cada altavoz con la señal que más le convenga. El filtrado se puede dividir en **pasivo** y **activo**. Los filtros pasivos van incluidos en las cajas acústicas con el altavoz y se alimentan con un único amplificador de potencia. En la Figura 2-5 se muestra el sistema pasivo de filtrado.

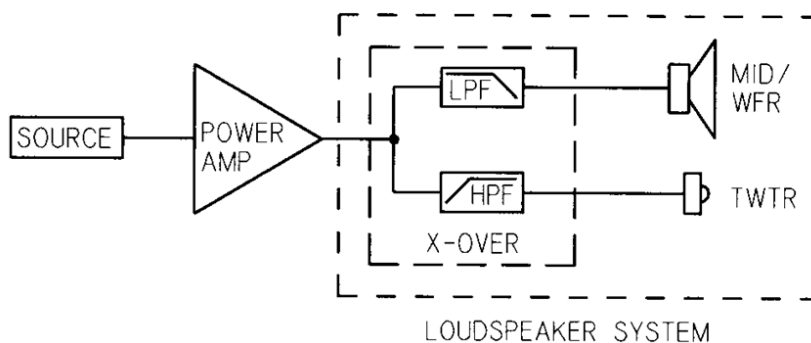


Figura 2-5. Sistema pasivo de filtrado.

Los filtros activos hacen la división del espectro de audio antes de llegar a los sistemas de amplificación. Por cada banda se necesita una amplificación independiente tal y como muestra Figura 2-6.

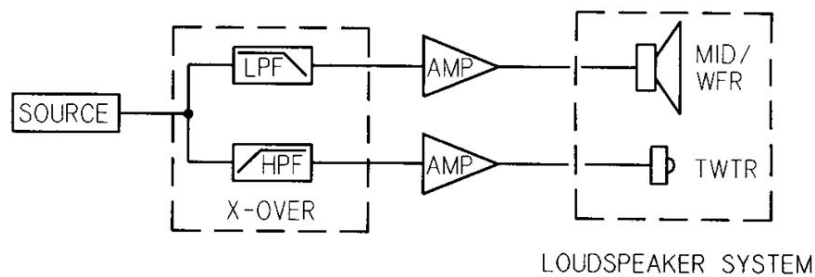


Figura 2-6. Sistema activo de filtrado.

2.1 Definición de altavoz

Un altavoz es un transductor electroacústico emisor. Está encargado de convertir la señal de audio en ondas de presión acústica. Dicha conversión se realiza en dos etapas representadas por el TEM y el TMA.

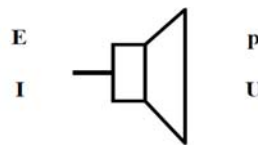


Figura 2-7. Altavoz.

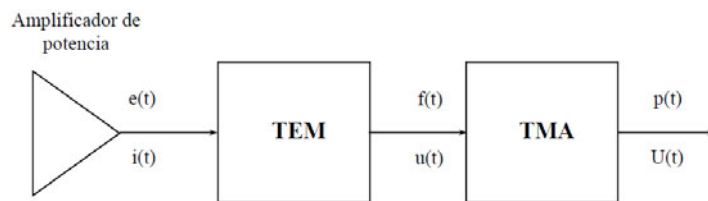


Figura 2-8. Esquema genérico de un altavoz conectado a un amplificador.

El TEM es el transductor electromecánico. Se encarga de transformar las señales eléctricas, tensión y corriente, en movimiento, representado por su fuerza y su velocidad. Es la etapa más importante del altavoz y a veces recibe el nombre de Motor o Driver.

El TMA es el transductor mecánico-acústico. Es el que transforma la vibración mecánica de un elemento en radiación acústica, representada por su presión y caudal. Esta parte recibe el nombre de **diafragma**, aunque en algunos altavoces la transducción se realiza con un adaptador denominado **bocina**.

Generalmente el diseño de los altavoces se centra en cumplir los requerimientos que se comentan a continuación:

- a- **Respuesta en frecuencia plana:** algo muy difícil de conseguir. Casi todos los altavoces tienen una respuesta en frecuencia muy irregular.
- b- **Respuesta en frecuencia amplia.** El margen de audio abarca aproximadamente 10 octavas, lo cual es muy difícil de cubrir con un solo altavoz. Por ello, se suelen utilizar sistemas de varios altavoces en los que cada uno se ocupa de un margen de frecuencias distinto.
- c- **Poca distorsión.** El diafragma es el elemento que más problemas da en este aspecto debido a su naturaleza mecánica.
- d- **Buena respuesta a transitorios,** lo cual significa, igual que en micrófonos, que responda rápidamente ante cambios bruscos de la señal eléctrica de excitación. Esta característica influye directamente en la respuesta en frecuencia del altavoz.

- e- **Eficacia y rendimiento.** Cuantifica el porcentaje de energía eléctrica que un altavoz transforma en radiación sonora. Esto interesa que sea lo más alto posible, ya que equivaldría a que el altavoz es capaz de producir un alto nivel de presión sonora con relativamente poca potencia eléctrica a la entrada. En la práctica se da que los altavoces son muy poco eficientes, generalmente solo aprovechan el 1% de la potencia eléctrica que se les proporciona, el resto se consume en la producción de calor.
- f- **Directividad.** Idealmente interesaría que un altavoz mantuviese la directividad con la frecuencia. De esa manera, los oyentes situados fuera del eje percibirían la misma respuesta que los que se sitúan en el. Es una condición muy difícil de conseguir con altavoces convencionales, en este aspecto funcionan mejor los sistemas basados en bocinas.

2.2 Clasificación de los altavoces

2.2.1 Según su TEM

En función del criterio que adoptaron sus diseñadores para construir los altavoces, podemos encontrarnos con los siguientes tipos:

- Altavoces dinámicos.
- Altavoces electrostáticos.
- Altavoces piezoeléctricos

a) Electrodinámicos, dinámicos o de conductor o de bobina móvil.

Se basan en la acción de la fuerza magnética que se ejerce sobre un conductor por el que circula una corriente variable, I - (la señal de audio), y que se mueve dentro del campo magnético constante, B , generado por un imán. El conductor puede ser una bobina que se mueve dentro del campo magnético fijo, producido por un imán anular o cilíndrico, en cuyo caso se tratarían de altavoces de bobina móvil. Los altavoces de bobina móvil son los más utilizados en la actualidad para cualquier tipo de aplicación. En la Figura 2-9 se muestra altavoz de bobina móvil.

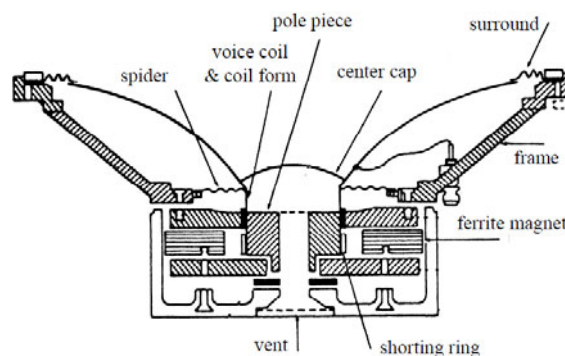


Figura 2-9. Altavoz de bobina móvil

Estos altavoces son muy robustos y permiten reproducir cualquier frecuencia. Se pueden construir altavoces donde en vez de una bobina haya un conductor plano sobre una cinta que se mueve dentro de un campo magnético constante. Este tipo de altavoces, raramente usados, se llaman de cinta que se muestra en la Figura 2-10.

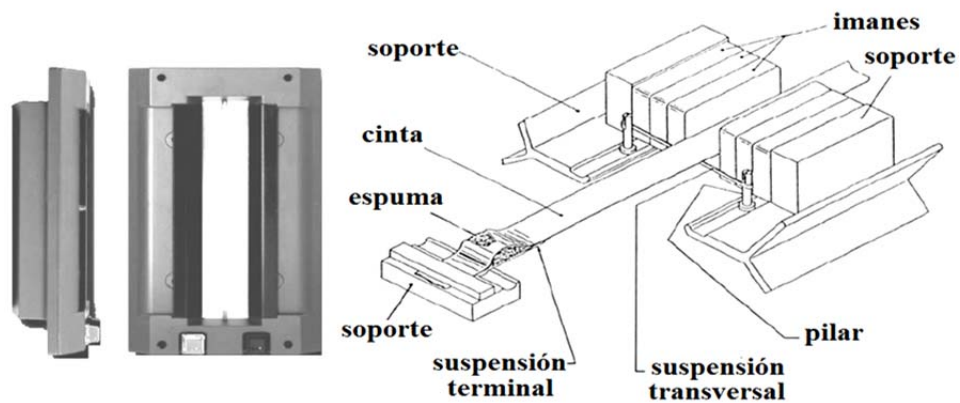


Figura 2-10. Altavoz dinámico de cinta.

Generalmente son utilizados para las vías de agudos en sistemas multi-vía. En ocasiones se utilizan para medios pero adoptan un tamaño bastante grande para radiar eficazmente en este margen de frecuencias.

Otro tipo de transductor dinámico es el altavoz isodinámico. Como se aprecia en la Figura 2-11, los imanes se disponen en forma de barra creando así un campo magnético lineal. El conductor móvil está unido al diafragma plano, se trata de una bobina también plana, dispuesta en zig-zag o en espiral, de tal manera que los sentidos de circulación de la corriente son contrarios en pistas paralelas. De esa forma la fuerza resultante sobre la bobina se aplica siempre en el mismo sentido para un determinado sentido de circulación de la corriente eléctrica.

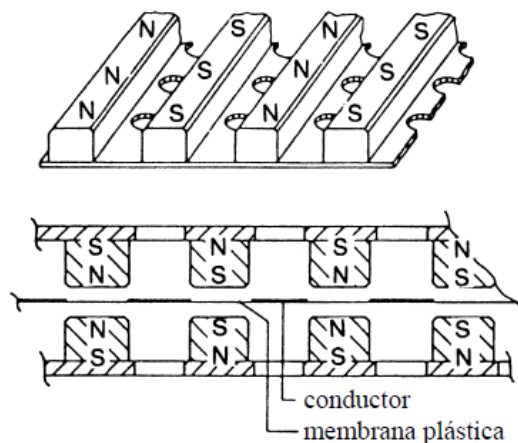


Figura 2-11. Transductor isodinámico. Disposición de los imanes, diafragma y conductor.

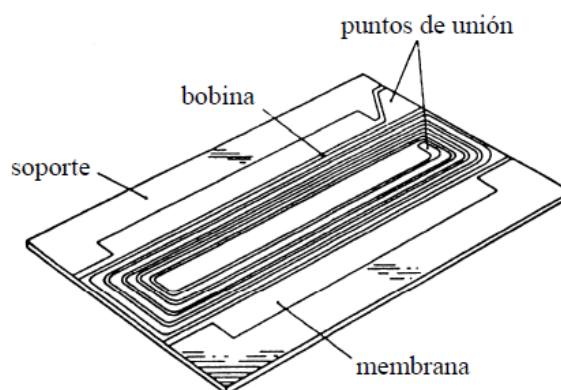
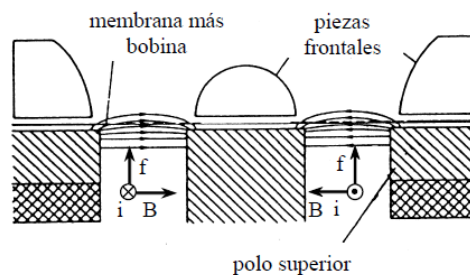


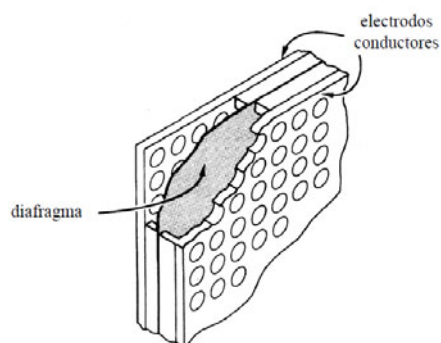
Figura 2-12. Disposición de la bobina plana.**Figura 2-13. Sentido del campo B, la fuerza $f(t)$ y la corriente $i(t)$ sobre pistas contiguas.**

Los altavoces isodinámicos tienen la ventaja de su reducido tamaño y espesor. Se usan sobre todo en la fabricación de auriculares de casco y de *tweeters*.

b) Altavoces electrostáticos o de condensador

Se basan en el mecanismo del condensador variable, usado también en los micrófonos de condensador. Su estructura consiste en tres placas metálicas que forman una especie de bocadillo. Las dos placas exteriores, rígidas y perforadas, constituyen el condensador. La placa interna, mucho más delgada, es el diafragma como se muestra en la Figura 2-14.

Cuando se aplica una señal eléctrica elevada a las dos placas que forman el condensador, las placas rígidas se mueven en función de este voltaje de entrada, generando un flujo magnético y el diafragma, situado en su interior, vibra.

**Figura 2-14. Detalle del altavoz electrostático. El diafragma suele ser de muy pequeño espesor y con alta resistividad superficial. Los electrodos fijos del condensador variable son de material conductor.**

La configuración por la que funcionan estos altavoces se denomina *push-pull*: una placa móvil entre dos placas fijas. Esta forma de operar proporciona mucha más linealidad que si funcionasen a tensión constante como los micrófonos de condensador. Son usados con mucha frecuencia en *tweeters*.

El principal inconveniente es su complicación electrónica: necesitan una fuente de tensión continua de alto voltaje y un transformador elevador de audio.

Proporcionan buenos medios y agudos, tienen un espesor muy pequeño y no necesitan caja. Su coste es bastante elevado.

c) Altavoces piezoeléctricos o cerámicos.

Están basados en el fenómeno de la piezoelectricidad, por el que ciertos materiales modifican sus dimensiones cuando se les somete a una tensión eléctrica aplicada entre dos superficies de dicho material. La tensión eléctrica es aplicada por medio de unos electrodos metálicos, colocados sobre un material piezoeléctrico sintético. En la Figura 2-15 se muestra la sección transversal de un altavoz piezoeléctrico.

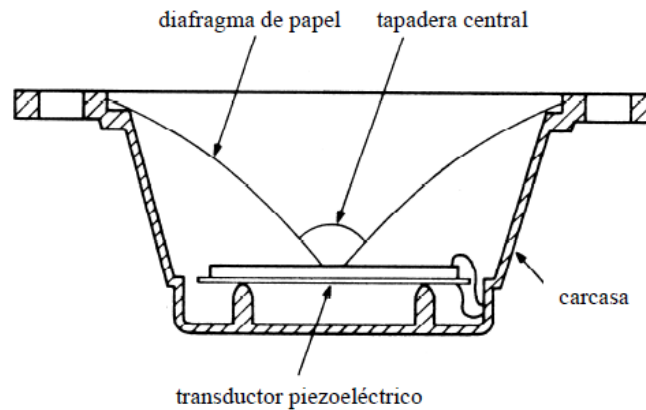


Figura 2-15. Corte esquemático de un altavoz piezoeléctrico lastrado por un diafragma de papel.

Para reducir la frecuencia de resonancia, que tiene la cerámica, y mejorar la radiación, estos altavoces suelen incorporar un diafragma. En el caso de la Figura 2-15 el diafragma es de papel. La configuración más utilizada en el diseño de los altavoces de cerámica es la denominada *bimorph*, que consiste en un par de cerámicas piezoeléctricas en contacto. Si la conexión eléctrica se hace de acuerdo con la dirección principal de polarización del material piezoeléctrico, cuando uno de los elementos del par se contraiga, el otro se alargará, con lo que la curvatura del *bimorph* es máxima y por tanto la radiación acústica también lo será. En la Figura 2-16 se muestran los dos tipos de configuración *bimorph*.

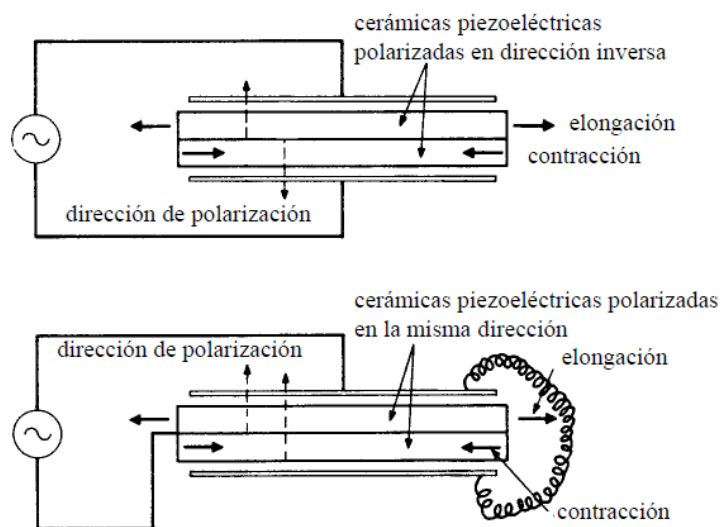


Figura 2-16. Configuración de un bimorph. (Arriba) Con las cerámicas se conectan eléctricamente en serie su dirección de polarización debe ser inversa. (Abajo) Cuando las cerámicas se conectan en paralelo su dirección de polarización debe ser la misma.

Son sencillos de fabricar, relativamente baratos, y funcionan bien a frecuencias altas.

d) Altavoces magnéticos o electromagnéticos.

Estos altavoces se basan en la Ley de Lenz, por la que un elemento móvil, en este caso el diafragma, vibra al variar el campo magnético que recorre el circuito magnético al que pertenece. La variación del campo magnético está producida por la señal de audio. Este tipo de transducción también es denominada 'reluctancia variable', ya que al vibrar el diafragma varía el equivalente magnético de la resistencia eléctrica, la reluctancia. A diferencia de los altavoces dinámicos el campo magnético en estos altavoces no es fijo, sino dependiente de la señal de audio. En la Figura 2-17 se muestran cortes transversales de dispositivos con este tipo de tecnología.

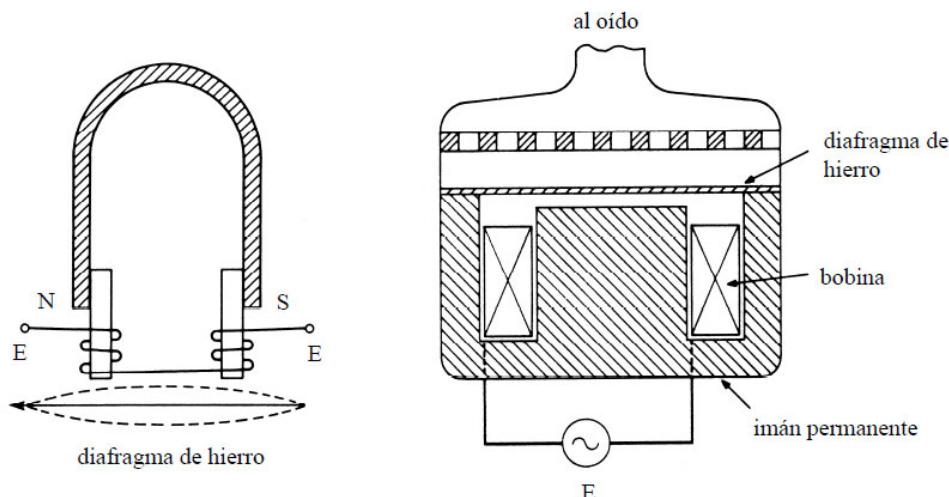


Figura 2-17. Diseño simplificado del auricular del teléfono de Bell (izqda.). Corte esquemático de un auricular de casco electromagnético (dcha.)

Son usados principalmente en auriculares de cascos y telefónicos.

El transductor de estos altavoces es alineal, de hecho, actúan como rectificadores mecánicos de la señal eléctrica de audio: tanto si la corriente de audio es positiva como negativa se produce la atracción del diafragma. Para que funcione adecuadamente se suele incluir un imán que proporcione un campo magnético estático.

Existen otros esquemas de radiación muy raramente utilizados como los emisores magnetostrictivos, usados solo para proyectores ultrasónicos, y neumáticos, usados para muy baja frecuencia. Hay altavoces denominados altavoces pasivos, que carecen de transductor electromecánico. Se trata de un diafragma impulsado por otro elemento mecánico o acústico y que forma parte de un sistema emisor más complejo. Los altavoces pasivos se suelen poner en los sistemas de refuerzo de graves de las cajas acústicas (*bass-reflex*), para mejorar la radiación del orificio o puerta que conecta la emisión posterior de un altavoz activo con el exterior. De esa forma, el altavoz pasivo cierra dicho orificio de *bass-reflex*, que de otra forma estaría abierto.

2.2.2 Según la banda de frecuencias reproducida

Dado que es muy complicado reproducir con un solo altavoz todo el margen de audio, existen altavoces especializados en radiar en determinadas zonas del espectro de audiofrecuencia.

- a) Altavoces de graves**, también conocidos como *woofers*, son los que normalmente reproducen frecuencias bajas, hasta los 500 Hz, dependiendo del caso. Los *subwoofers* están especializados en la radiación por debajo de 100Hz.
- b) Altavoces de medios**, también denominados *squeakers* o *midrangers*. Reproducen la banda media de frecuencias, desde los 500 hasta los 5kHz.
- c) Altavoces de agudos**, o *tweeters*, están especializados en radiar frecuencias por encima de unos 5 kHz.

Los márgenes de frecuencia anteriores varían con la naturaleza del sistema de radiación multivía, dependerá de si el espectro se divide en dos, tres, cuatro o más vías. Los mal llamados altavoces de banda ancha, en realidad se tratan de altavoces de medios con una margen de frecuencias mayor que uno de medios convencional, pero no reproducen con fiabilidad todas las frecuencias audibles. Se usan sobretudo en electrónica de consumo, como televisores, radios, automóviles, etc.

2.2.3 Según el TMA

a) **Altavoces de cono**, o de radiación directa, son aquellos en los que el diafragma del altavoz se enfrenta directamente al aire. La mayoría de altavoces comerciales son de este tipo.

b) **Altavoces de bocina**, o de radiación indirecta, son aquellos que aunque existe diafragma, este no radia directamente al aire sino que lo hace a través de una bocina. La bocina es un tubo rígido de abertura progresiva que sirve para adaptar la radiación desde un orificio muy pequeño al aire. Este orificio es la salida de caudal en los motores de compresión, que es como se llama el altavoz en este tipo de sistemas.

Los altavoces de bocina se caracterizan por una mayor radiación que los altavoces de cono, lo que conlleva una mayor sensibilidad y eficiencia. Además en su diseño se puede tener control sobre la directividad del sistema. Solo radian en medias y altas frecuencias, aunque se pueden diseñar altavoces de cono con bocinas especiales para baja frecuencia.

2.3 Altavoz dinámico de bobina móvil

Aunque se han explicado diversas formas de transducción, en la práctica la mayoría de sistemas comerciales que se fabrican están constituidos por altavoces dinámicos de bobina móvil. En la Figura 2-18 se muestra el corte transversal de un altavoz de bobina móvil.

2.3.1 Características constructivas de un altavoz de bobina móvil.

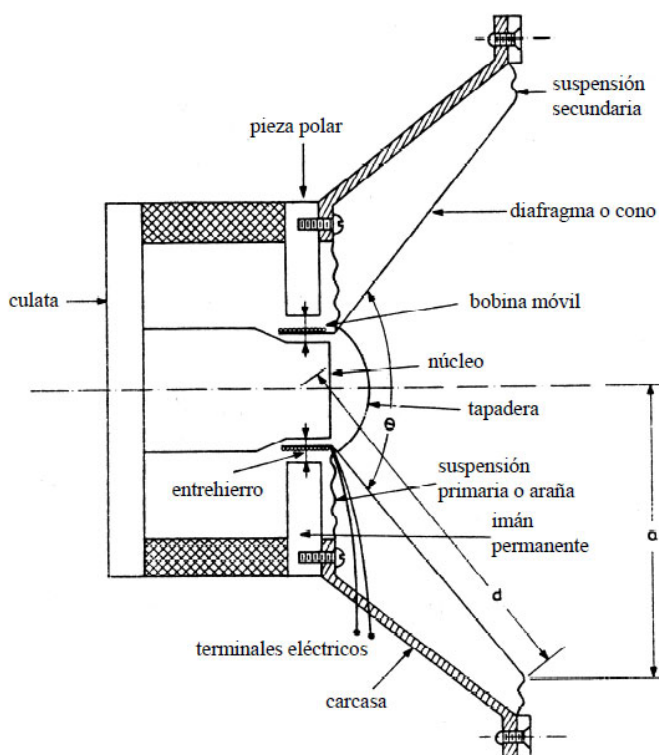


Figura 2-18. Corte esquemático de un altavoz de bobina móvil.

a) Diafragma o cono

Constituye el TMA. Es la parte del mismo que se enfrenta al medio fluido, el aire, y es el que, al vibrar, produce ondas acústicas de presión y caudal. El diafragma es una de las partes más importantes del altavoz, ya que de él dependen las propiedades más importantes del mismo. En la Figura 2-19 se muestran todas las partes que conforman un altavoz de bobina móvil.

-Tamaño del diafragma

El tamaño del diafragma influye en la eficiencia, directividad y respuesta en frecuencia.

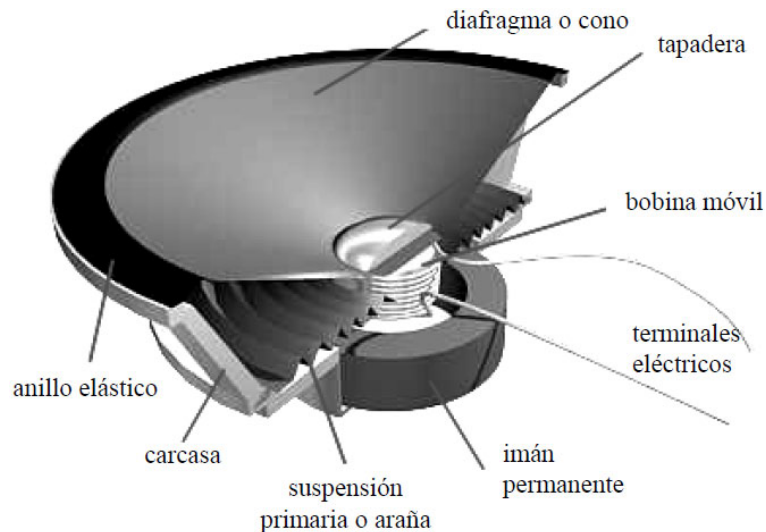


Figura 2-19. Corte en perspectiva de un altavoz dinámico de bobina móvil.

La respuesta en frecuencia de un altavoz dinámico, es su parte útil, está dominada por la masa. Es decir, su respuesta en presión, y por tanto la sensibilidad, es inversamente proporcional a la masa, al igual que su eficiencia y rendimiento. Por ello los altavoces con mayor diafragma deben tener un factor de fuerza, **Bl**, mayor, que normalmente se traduce en una mayor imán y circuito magnético.

El ancho de banda también está determinado por el tamaño del diafragma y por la masa del mismo. La frecuencia límite inferior en la respuesta en frecuencia de un altavoz dinámico está determinada por la frecuencia de resonancia mecánica, que es inversamente proporcional a la raíz cuadrada de la masa mecánica del diafragma. De ahí que los diafragmas grandes (mayor masa) respondan mejor a las bajas frecuencias. También se puede hacer que un diafragma pequeño radie en graves, solo hace falta que este sea más blando o elástico. El problema principal viene en el comportamiento no lineal de un pequeño diafragma radiando frecuencias bajas debido a las limitaciones mecánicas.

Lo que sí es cierto es que un diafragma grande no puede radiar en alta frecuencia de forma eficiente, debido a su poca rigidez. Luego los altavoces de agudos son pequeños. La directividad de un radiador es proporcional al factor ka , con $k=\omega/c$ el número de onda y a el radio del diafragma. Luego un altavoz será más directivo conforme aumente su frecuencia.

-Forma del diafragma

Cúpulas o domos: Utilizadas sobre todo para unidades de alta frecuencia debido a la rigidez proporcionada por su curva. Existen cúpulas cóncavas y convexas como se ve en la Figura 2-20.

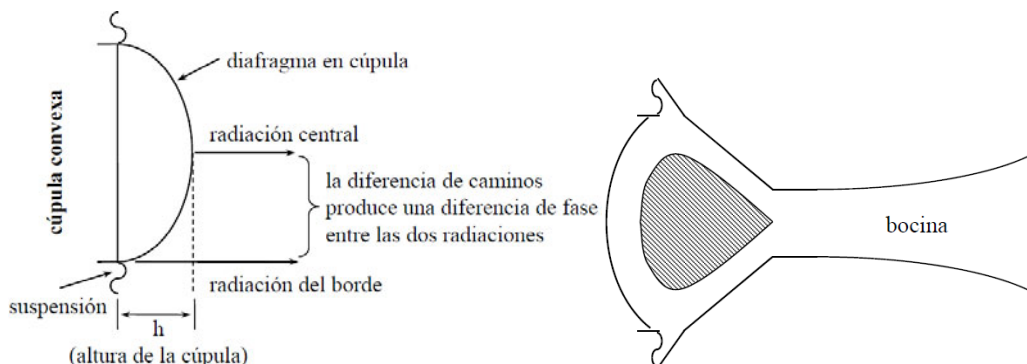


Figura 2-20. Geometría de una cúpula convexa (izqda.) empleada como diafragma de un altavoz. Cúpula cóncava (dcha.).

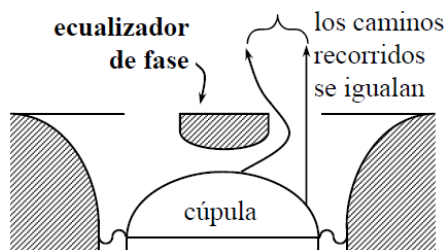


Figura 2-21. Detalle de un ecualizador de fase.

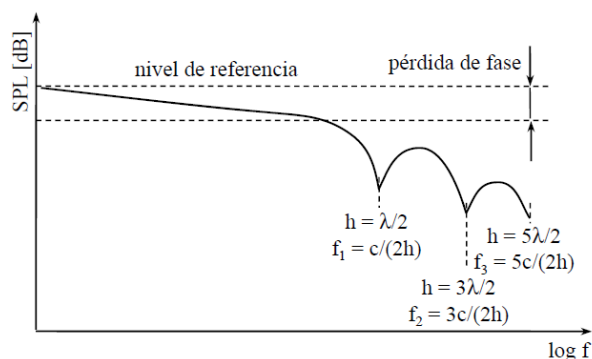


Figura 2-22. Respuesta en frecuencia típica de un altavoz de cúpula debida a la pérdida de fase.

Las pérdidas de fase, Figura 2-22, se producen por la diferencia de caminos que hay entre la radiación central y la radiación del borde del diafragma. Para corregir este defecto hay que reducir el tamaño de la cúpula, h . Pero por otra parte la profundidad favorece a una directividad baja. Entonces para aminorar la pérdida de fase sin hacer más directiva la respuesta se introducen los llamados ecualizadores de fase, Figura 2-21, elemento que pretende igualar el trayecto recorrido por la onda acústica procedente de la parte central del diafragma, con respecto al recorrido por la radiación lateral.

Conos: Los diafragmas de cono son los más comunes en la construcción de un altavoz dinámico. Esta configuración permite una gran superficie de radiación con una pequeña superficie impulsora. Además consigue una mayor sujeción de la parte móvil del altavoz, evitando el descentrado de la bobina, por lo que se usan en *woofers* y *subwoofers*, en los que el diafragma tiene grandes desplazamientos.

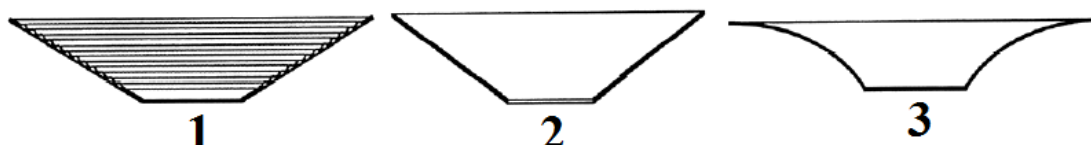


Figura 2-23. Diversos perfiles de diafragmas cónicos en un altavoz dinámico de bobina móvil. (1) Perfil recto con escalones circulares. (2) Perfil recto liso. (3) Perfil curvo.

Diafragmas para altavoces de banda ancha. Aunque es muy difícil que un solo altavoz sirva para cubrir toda la banda de audio, existen configuraciones de diafragmas que facilitan la efectividad del mismo en un amplio rango de frecuencias.

i-Diafragmas elípticos. Son ligeramente más rígidos que los circulares. Consiguen tener una superficie grande y un radio pequeño a la vez. Con la superficie grande se mejora la radiación en graves y con el radio pequeño evita las vibraciones parciales del diafragma. En la Figura 2-24 se muestra un diafragma elíptico.

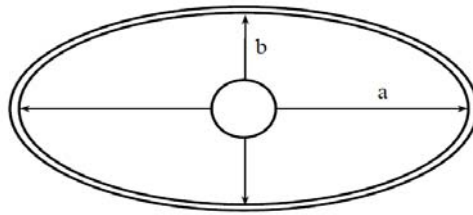


Figura 2-24. Diafragma elíptico.

ii-Altavoces de doble cono. Son una alternativa más económica que los de dos vías. Consisten en dos diafragmas unidos por una suspensión mecánica tal y como se muestra en la Figura 2-25.

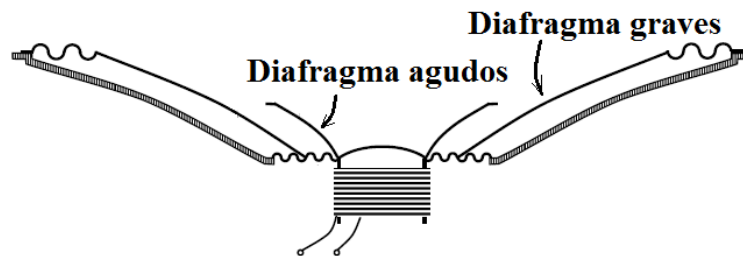


Figura 2-25. Estructura de los altavoces de doble cono.

En frecuencias bajas, el diafragma de graves es el que radia, mientras que en altas frecuencias, la radiación se hace por el diafragma de agudos. No hay que confundirlos con los altavoces coaxiales que tienen bobinas distintas para cada diafragma.

iii-Altavoces coaxiales. Tal y como se muestra en la Figura 2-26 están constituidos por dos diafragmas totalmente aislados mecánicamente pero que comparten el mismo eje de radiación. Los dos altavoces están unidos por la parte eléctrica a través de filtros de cruce y ambos tienen el circuito magnético completo, con sus bobinas, imanes y piezas polares correspondientes. Por ello, al tratarse de un sistema de dos vías necesitan un filtro de cruce que divida el espectro de audio. Tienen la ventaja de que ahorran mucho espacio y su radiación tiene una simetría cilíndrica perfecta. Pero es muy complicado diseñar altavoces coaxiales con una respuesta en graves como la de un *woofer* y en agudos como la de un *tweeter*. Por ello en sistemas profesionales se prefieren altavoces de dos vías completamente independientes.

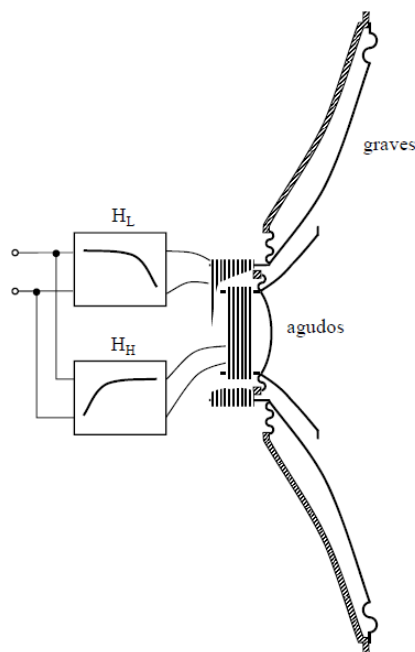


Figura 2-26. Estructura de un altavoz coaxial.

b) Bobina móvil.

Es el conductor encargado de transformar las variaciones de tensión eléctrica de sus extremos en fuerza magnética y por tanto en vibración mecánica y radiación acústica.

Normalmente está fabricada de aluminio anodizado y cobre, mecanizados en hilos esmaltados para evitar el contacto eléctrico entre las diferentes espiras de la bobina. Puede ser o cilíndrica o cuadrada, y se enrolla en 2, 4 o 6 capas, dependiendo esto último de la impedancia eléctrica necesaria, de la potencia eléctrica y la sensibilidad requerida. Para que la propia bobina no apantalle el campo magnético con sus diferentes capas, a menudo se utiliza una única capa de hilo plano en forma de cinta (*ribbon*). En la Figura 2-27 se muestra un corte transversal de la bobina móvil.

El **soporte** que sustenta la bobina es un cilindro hueco fabricado de cartón, aluminio o plástico denominado *former*. La bobina se une al *former* con cemento termo-resistente, tal y como se muestra en la Figura 2-28, aunque actualmente se emplean bobinas auto-soportadas que no lo necesitan.

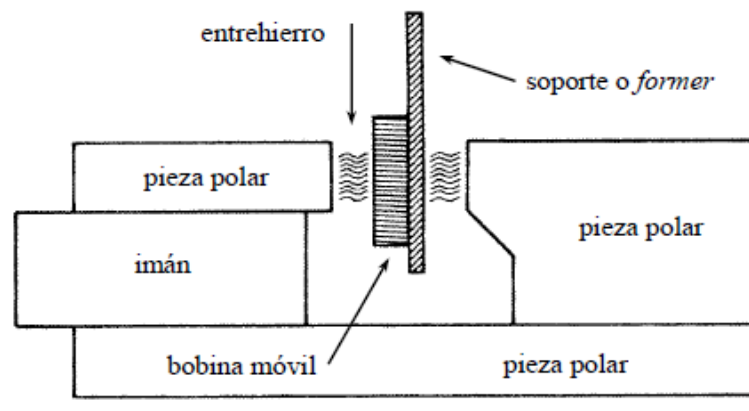


Figura 2-27. Corte transversal del altavoz de bobina móvil.

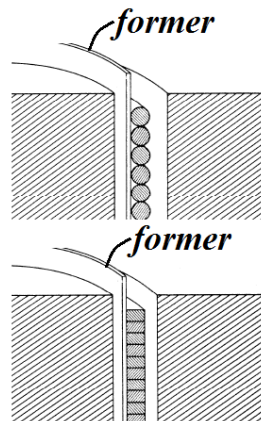


Figura 2-28. Bobina cuadrada y circular dispuestas en la cara externa del former.

La potencia eléctrica y la disipación de calor tienen gran dependencia con las dimensiones y naturaleza de la bobina móvil y el *former*.

Las temperaturas alcanzadas por la bobina cuando se trabaja a plena potencia pueden ser de en torno a los 200 °C, luego el *former* y la bobina han de hacerse con materiales que soporten dichas temperaturas. Además se han de usar técnicas adicionales de disipación de calor como hacer *formers* de gran diámetro, el uso de ferrofluidos (solución viscosa que conduce mejor el calor que el aire y lo traspasa a la estructura del altavoz) y fabricar los *formers* de materiales metálicos no magnéticos. En baja frecuencia, debido a los grandes desplazamientos del

diafragma, se producen corrientes de aire, que bien aprovechadas suponen una solución más para la disipación del calor.

Hay una estrecha relación entre el tamaño del *former* y la potencia máxima que soporta un altavoz: a mayor diámetro mayor potencia ya que se puede disipar más calor. Doblar el diámetro del *former* puede suponer doblar la potencia máxima del altavoz. El aumento del *former* implica el aumento del diafragma del altavoz. La Tabla 2-1 muestra la relación de tamaños entre el *former*-diafragma.

Tabla 2-1. Diámetros de altavoz más adecuados dependiendo de los diferentes diámetros de bobinas existentes.

diámetro de la bobina (<i>former</i>) [mm]	diámetro del altavoz [mm]
20 – 25	100 – 170
25 – 37	200 – 250
44 – 100	250 – 380

c) Imán y circuito magnético.

En un altavoz dinámico el imán permanente es el elemento que se encarga de producir el campo magnético representado por la inducción B .

Los imanes suelen ser metálicos o de cerámica. Actualmente se usan más frecuentemente imanes cerámicos, constituidos por partículas (ferritas) de óxidos ferromagnéticos depositadas sobre soportes de naturaleza cerámica. Para orientar a las partículas, el material, cuando aún está caliente, es sometido a un intenso campo magnético. Aunque los imanes cerámicos son más ligeros que los metálicos también son más frágiles. Además no sufren el proceso de desmagnetización como los imanes metálicos, son temporalmente más estables.

La forma del imán puede ser cilíndrica o anular, aunque la anular es la más usada ya que produce menos pérdidas de campo magnético en la bobina. En la Figura 2-29 se muestran imanes de distintas formas.

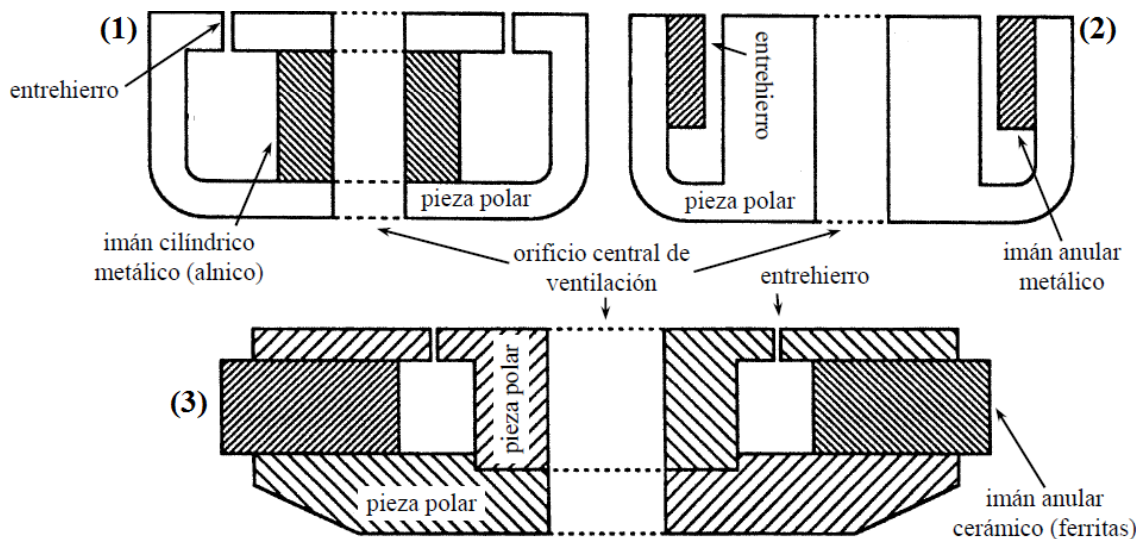


Figura 2-29. Imán cilíndrico metálico con orificio central de ventilación (1). Imán metálico en anillo, enfrentado directamente al entrehierro (2). Imán anular cerámico de ferrita.

El **entrehierro** es el espacio donde se ubica la bobina e interesa que este sea estrecho y de gran superficie para proporcionar un campo magnético intenso. Es el que proporciona la mayoría de la reluctancia magnética del circuito magnético (hay que recordar que el campo magnético en el entrehierro es inversamente proporcional a dicha reluctancia). La reluctancia viene dada por la ecuación (2-1):

$$R = \frac{l_0}{\mu_0 S_0} \text{ A/Weber. (2-1)}$$

Con l_0 y S_0 la longitud y la superficie del entrehierro respectivamente y μ_0 la permeabilidad magnética del aire.

El entrehierro también ayuda a la disipación de calor.

Las **piezas polares** son los elementos metálicos que se encargan de dirigir y concentrar el campo magnético B desde el imán hasta la bobina, de forma que las líneas de campo corten perpendicularmente a las espiras de la bobina.

La configuración geométrica del circuito magnético y la relación del tamaño de la bobina con la altura del entrehierro influyen en la distorsión del altavoz en baja frecuencia. En baja frecuencia puede que la bobina llegue a salirse del entrehierro, lo que provocaría que el campo magnético al que está sometida no sea constante y producirse así un comportamiento no lineal de la respuesta del altavoz tal y como se muestra en la Figura 2-30.

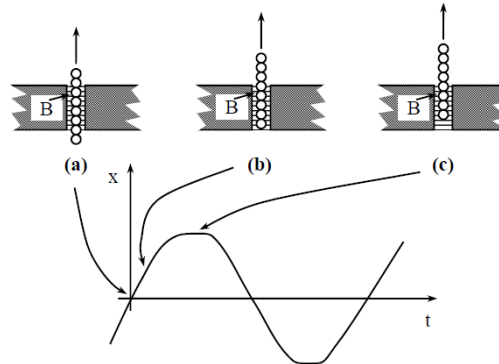


Figura 2-30. Distorsión en baja frecuencia debido a la salida de la bobina del entrehierro. Mientras que en (a) y (b) la bobina se mantiene dentro del entrehierro el comportamiento del altavoz es lineal, cuando en (c) hay parte de la bobina que sale del entrehierro se produce un corte en la respuesta.

Para solucionar este problema se pueden hacer bobinas mayores que la altura del entrehierro o hacerlas menores, como se puede ver en la Figura 2-31-a y b. Si se hacen bobinas mayores que la altura del entrehierro el número de bobinas sometidas al entrehierro siempre será constante. Los inconvenientes son un desaprovechamiento de bobina, aumento de la impedancia eléctrica, aumento de la masa mecánica del altavoz y disminución de la sensibilidad. Si se opta por hacer la bobina menor que la altura del entrehierro, ésta siempre estará sometida a la inducción magnética. La impedancia eléctrica es relativamente baja, pero se desaprovecha potencia del imán, perdiéndose así sensibilidad.

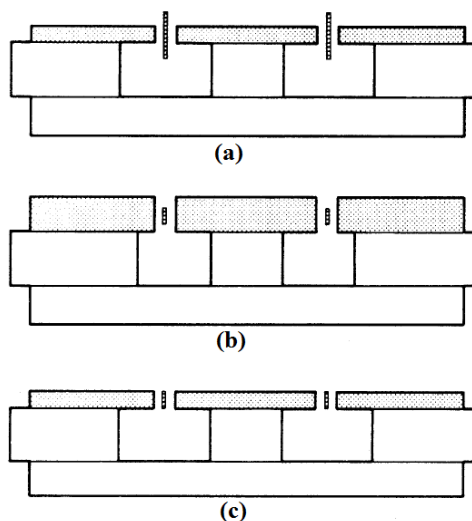


Figura 2-31. Bobina mayor a la altura del entrehierro (a). Bobina menor que el entrehierro (b). Bobina igual que el entrehierro (c).

La configuración que se muestra en la Figura 2-31-c apenas se usa, ya que un mínimo desplazamiento del diafragma produciría distorsión. Generalmente la relación entre bobina-altura entrehierro suele ser de 1/3 de bobina dentro del campo magnético.

d) Suspensión primaria o araña.

Es el principal elemento responsable de dar elasticidad al movimiento del diafragma y por tanto está relacionada directamente con la compliancia mecánica del mismo.

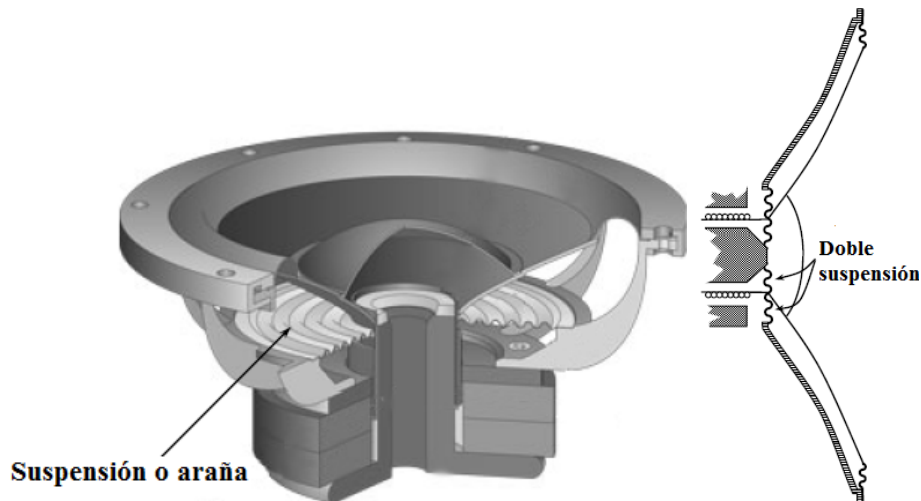


Figura 2-32. Suspensión primaria o araña (izqda.). Doble suspensión en un woofer (dcha.)

Además, la suspensión primaria permite el centrado y estabilización del diafragma, impidiendo que este se mueva en otra dirección que no sea paralela al eje del altavoz. La suspensión está hecha de material flexible con corrugaciones o no dependiendo de la excursión máxima esperada para el diafragma. Generalmente los altavoces de bajas frecuencias si tienen suspensión con corrugaciones.

Materiales típicos de fabricación de la araña son el nylon y el poliéster, recubiertos siempre de resina epoxy para darles forma.

La suspensión mecánica puede ser dura o blanda. Los altavoces 'duros' se suelen usar en aplicaciones profesionales donde es necesario prevenir un excesivo desplazamiento del diafragma. Un altavoz 'blando' está asociado habitualmente a una elevada distorsión para bajas fuerzas de excitación y son usados en instalaciones domésticas o de baja potencia.

e) Otros elementos

Suspensión secundaria: o superior, es la encargada de sujetar la parte superior del diafragma a la estructura del altavoz. Se usa generalmente en unidades de graves. Se fabrican de cartón, goma y plástico.

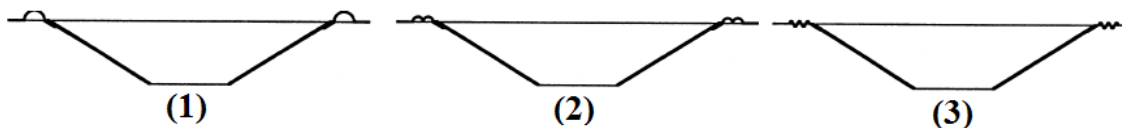


Figura 2-33. Suspensión secundaria de media vuelta (1), de doble media vuelta (2) y tipo acordeón (3).

Sirve sobre todo para centrar el diafragma, impidiendo que éste cabecee cuando hay grandes vibraciones del mismo. Además sirve de terminación resistiva del diafragma para evitar posibles modos propios transversales en diafragmas muy grandes.

Tapadera o cúpula: es la encargada de sellar la parte central del altavoz evitando así la entrada de polvo en su interior. También sirve para aminorar la amplitud de los modos propios transversales del diafragma y para subir su frecuencia.



Figura 2-34. Tapadera.

Carcasa: (*frame* o *basket*) es la estructura que soporta el altavoz. Tiene que ser rígida y pesada. Suele ser de acero galvanizado o aluminio, materiales no magnético para que no deformen el campo magnético interno del altavoz. No debe impedir el flujo de aire.

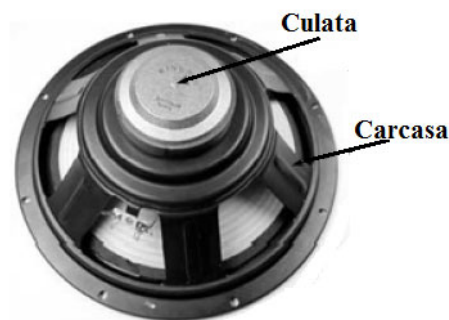


Figura 2-35. Detalle de la carcasa y culata de un altavoz de graves.

Yugo o culata: es la tapadera posterior del altavoz. Está realizada de material ferromagnético, y es la encargada de cerrar el circuito magnético. Además sella el altavoz por la parte trasera a no ser que tenga un agujero de ventilación (en unidades de graves muy potentes).

2.4 Especificaciones de un altavoz

- Impedancia.
- Potencia.
- Sensibilidad.
- Respuesta en frecuencia.
- Directividad. Factor Q.
- Eficiencia.
- Distorsión.

Las medidas de un altavoz dependen del montaje electroacústico del mismo y del medio en el que se encuentre el transductor. Las medidas se deben realizar en campo libre, es decir en cámara anecoica o semianecoica.

Cuando las especificaciones electroacústicas se refieren a un altavoz individual, sin estar montado en una caja, se ha de medir en pantalla infinita. Como la pantalla infinita es inalcanzable la norma IEC proporciona unas configuraciones de baffle estándar cuyo plano estructural se es el que se muestra en la Figura 2-36. Las dimensiones se detallan en la Tabla 2-2.

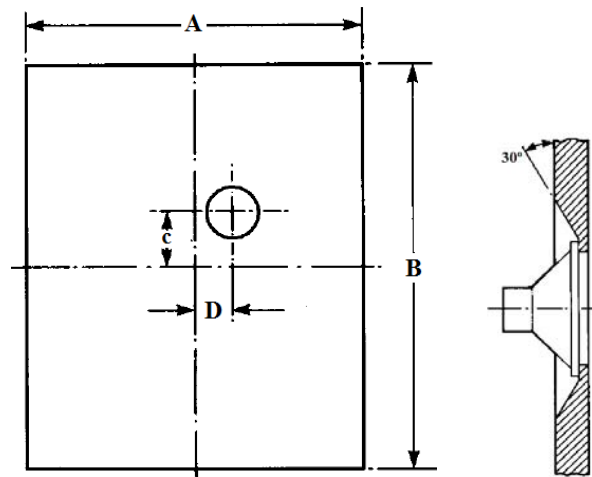


Figura 2-36. Baffle estándar recomendado por la norma IEC.

Tabla 2-2. Tamaños recomendados del baffle de medida (norma IEC).

tamaño nominal del altavoz [mm]	dimensiones del baffle [mm]			
	A	B	C	D
200	1350	1650	225	150
250	1690	2065	280	190
315	2025	2475	340	225
400	2530	3090	420	280
500	3040	3715	505	340

2.4.1 Impedancia

La impedancia eléctrica de entrada, $Z_{ee} [\Omega]$, es la relación compleja entre tensión eléctrica, E, aplicada y la corriente, I, absorbida por un altavoz. Se suele representar el módulo $|Z_{ee}(f)| [\Omega]$ en función de la frecuencia, aunque algunos fabricantes especifican también la parte real, $|Ree(f)| [\Omega]$, y la parte reactiva, $|X_{ee}(f)| [\Omega]$, en función de la frecuencia. También se proporciona la resistencia en continua R_e . En la Figura 2-37 se muestra como varía la impedancia de un altavoz en función de la frecuencia.

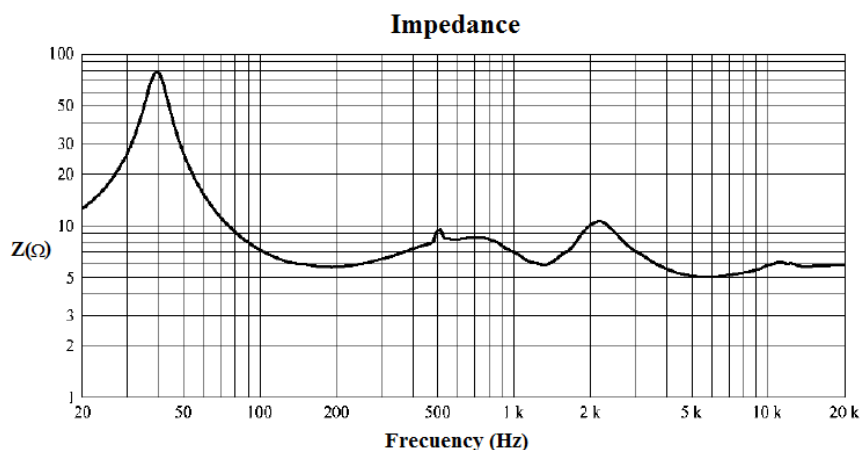


Figura 2-37. Impedancia eléctrica de entrada, $|Z_{ee}(f)| [\Omega]$, de un altavoz coaxial de doble vía.

Cuanto menor sea la impedancia eléctrica que 've' el amplificador de potencia mayor será la corriente que tendrá que entregar y mayor el calentamiento de su etapa de salida, con la posibilidad de que éste se queme. Interesa que la impedancia eléctrica de entrada de un altavoz no descienda demasiado debajo de la impedancia nominal del mismo.

La **impedancia nominal**, Z_{nom} [Ω], es la resistencia eléctrica a la que se puede aproximar el altavoz y es la que sustituye a éste en los cálculos de potencia eléctrica, tensión y corriente. La expresión que relaciona potencia eléctrica absorbida por el altavoz y tensión eléctrica aplicada se calcula con la ecuación (2-2):

$$P_e = \frac{E^2}{Z_{nom}} W. \quad (2-2)$$

Los valores típicos son 4 y 8 Ω . Cuando conectemos un altavoz a un equipo debemos tener en cuenta este valor, ya que tenemos que conservar su impedancia de salida. Por ejemplo, si tenemos un amplificador con una impedancia de salida de 8 ohmios, podemos situar 1 altavoz de 8 Ω , 2 altavoces de 4 Ω en serie, 2 altavoces de 16 Ω en paralelo, etc.

Así, una señal eléctrica de 1 W de potencia desarrollada sobre $Z_{nom} = 8 \Omega$ equivale a una tensión eléctrica de $E = \sqrt{1 \cdot 8} = 2,83$ V. La tensión es la RMS o eficaz.

Se debe cumplir que $|Z_{ee}| \geq 0,8 Z_{nom}$ para que el amplificador de potencia no sufra ningún daño.

2.4.2 Potencia

La **potencia máxima** en altavoces es una indicación de seguridad eléctrica ante la ruptura, por eso se preferirá altavoces que trabajen muy por debajo de su potencia eléctrica máxima. La definición de potencia eléctrica máxima depende del tipo de señal eléctrica con la que se excite al altavoz y de la duración de la prueba. El organismo de normalización IEC especifica diversas normas sobre potencia máxima en altavoces. El IEC junto con el AES (*Audio Engineering Society*) son las normas más seguidas por los fabricantes.

La **potencia nominal de ruido**, equivalente a la potencia de programa, musical o de pico, es la potencia eléctrica que aguanta un altavoz o sistema de altavoces sin sufrir daño permanente cuando se le excita durante un tiempo elevado con una señal de banda ancha filtrada. La norma IEC establece un tiempo de prueba de al menos 100 horas y la señal será un ruido rosa filtrado según la curva de la Figura 2-38.

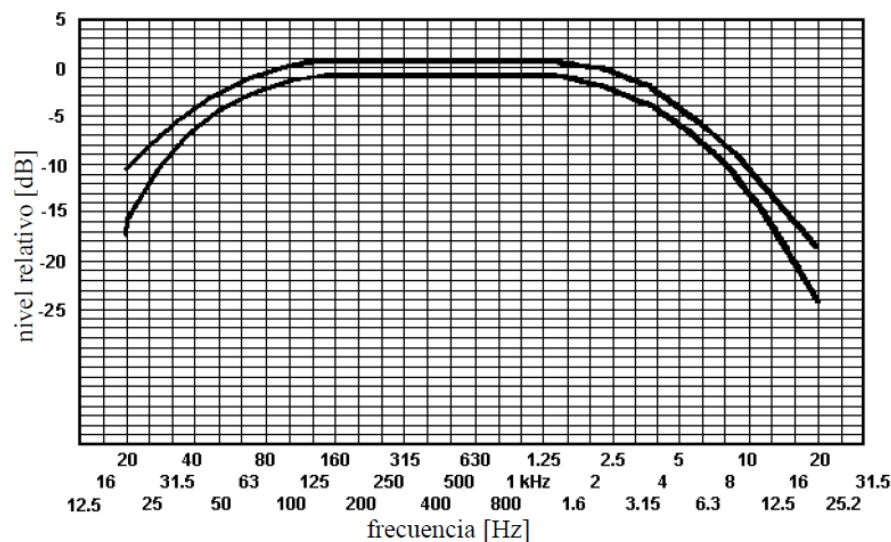


Figura 2-38. Función de transferencia (gálbo de tolerancia) para el filtrado de ruido según la norma IEC.

La norma AES recomienda al menos 2 horas con ruido rosa, pero filtrado en una década, dependiendo de la banda de trabajo del altavoz. En sistemas de tres vías se filtra una década, en los de dos vías 1,5 décadas.

Luego, según lo indicado la norma IEC, ésta es más adecuada para sistemas multivía que cubran todo el espectro de audio y la norma AES para medir la potencia de cada altavoz del sistema.

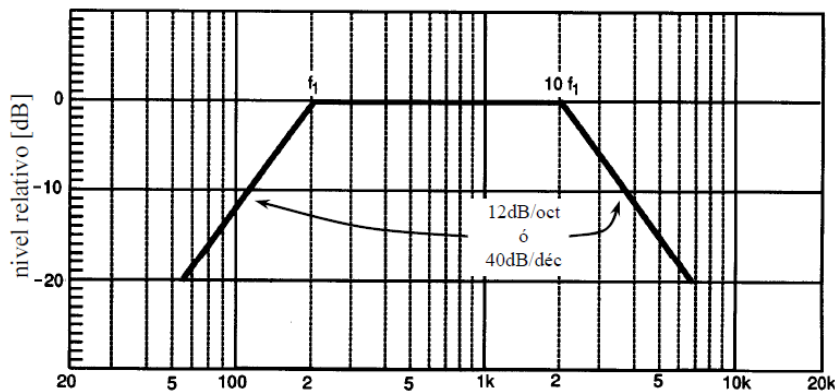


Figura 2-39. Función de transferencia para el filtrado de ruido según AES.

La **potencia nominal sinusoidal o continua** es la máxima potencia que soporta un altavoz o sistema de altavoces durante un elevado periodo de tiempo cuando se le excita con un barrido de tonos cíclico en una determinada banda de trabajo, acorde con la gama de frecuencias que es capaz de reproducir adecuadamente el altavoz. A la potencia máxima determinada por una señal sinusoidal se le suele denominar **PHC** (*Power Handling Capacity*). La IEC determina una duración de la prueba de 100 horas.

Otra consideración es la **potencia del amplificador**. En amplificadores, la potencia máxima se especifica para el recorte de la onda de salida. La potencia nominal se mide con tonos puros y la potencia musical con señales de tipo impulsivo, parecidas al ruido rosa. La potencia musical suele ser de 1 a 3 dB superior que la nominal. Para que el altavoz esté totalmente protegido es necesario que la potencia nominal del amplificador sea igual o menor que la potencia de programa del altavoz.

En un sistema amplificador altavoz, el que gobierna la potencia entregada, para una Z_{nom} fija del altavoz, es el amplificador. Es decir, por mucha potencia que tenga el altavoz, si el altavoz es pequeño, la potencia disipada por el altavoz será la que entregue el amplificador.

Por último cabe mencionar también de la **potencia de servicio**, como aquella necesaria para obtener un nivel determinado a una determinada distancia; o la **potencia máxima antes de distorsión**, como aquella necesaria para que el altavoz alcance una distorsión armónica determinada (THD ó TH2), normalmente del 3%.

2.4.3 Sensibilidad

La sensibilidad, S , de un altavoz es el nivel de presión sonora medido en el eje y a 1 metro de distancia, en cámara anecoica, cuando se le excita con ruido rosa filtrado en su banda útil y con una potencia de $P_e = 1$ W. Se especifica en dB. A diferencia de los micrófonos los altavoces tienen una respuesta en frecuencia muy coloreada, con lo que la sensibilidad no puede medirse con un tono puro. La sensibilidad informa sobre la sensación global de nivel que ofrecerá el altavoz cuando trabaje con señal de banda ancha. Si en la banda de trabajo el altavoz contiene algún pico o agujero, la sensibilidad medida será mayor o menor respectivamente.

Aunque la sensibilidad se especifica a 1m, se recomienda hacer la medida con una distancia mayor, por ejemplo 2 metros, debiendo corregirse el valor obtenido para adecuarlo a la referencia de 1 m. Con esto se evitan ambigüedades e imprecisiones, sobre todo cuando medimos un altavoz con mucha profundidad.

Aceptando que todos los altavoces son lineales para excitaciones próximas a la sensibilidad, y que se cumple la ley de divergencia esférica por estar en cámara anecoica, la sensibilidad S se puede calcular con la ecuación (2-3):

$$S = \text{SPL}(1 \text{ W}, 1 \text{ m}) = \text{SPL}(P_e, r) - 10\log(P_e) + 20\log(r) \text{ dB. (2-3)}$$

con P_e la potencia eléctrica aplicada al altavoz en su banda de trabajo y r la distancia usada en la prueba.

El rango de valores de sensibilidad característico para altavoces comerciales es muy amplio, desde 80dB_{SPL} hasta unos 115dB_{SPL}.

2.4.4 Respuesta en frecuencia

La respuesta en frecuencia de un altavoz está representada por su curva de nivel en función de la frecuencia y medida en campo libre (cámara anecoica). Se puede hacer por barrido de tonos o en tercios de octava. Todo a 1 W. Pero actualmente con los analizadores espectrales modernos (FFT, *Fast Fourier Transform* o CPB, *Constant Percentage Band*) no se suelen usar los dos métodos anteriores. Se usa un ruido de banda ancha y el espectro captado por un micrófono. Con esto se obtienen niveles muy bajos ya que, mientras que en el barrido de tonos se usaba 1W para una frecuencia determinada, cuando excitamos con un ruido de banda ancha el watio se reparte en toda esa franja del espectro. Para evitar este problema se puede medir la función de transferencia. En la Figura 2-40, Figura 2-41y en la Figura 2-42 se muestran distintas respuestas en frecuencia según su método de obtención.

La **función de transferencia** de un altavoz, considerada a una distancia r , es la relación entre la presión radiada y la tensión aplicada y evaluada para cada una de las frecuencias del espectro de audio. La función de transferencia viene dada por la ecuación (2-4):

$$H(r)[dB] = 20 \log \frac{p(r)}{E} \text{ dB. (2-4)}$$

Normalmente conviene convertir la curva de función de transferencia H en curva de nivel de presión sonora, SPL, de tipo barrido de tonos mediante la ecuación (2-5) :

$$\text{SPL}(r, E) = 20 \log \frac{p(r)}{p_{\text{ref}}} = 20 \log \frac{p(r)}{E} + 20 \log E - 20 \log p_{\text{ref}} = H(r)[dB] + 20 \log E - 20 \log p_{\text{ref}} \text{ dB. (2-5)}$$

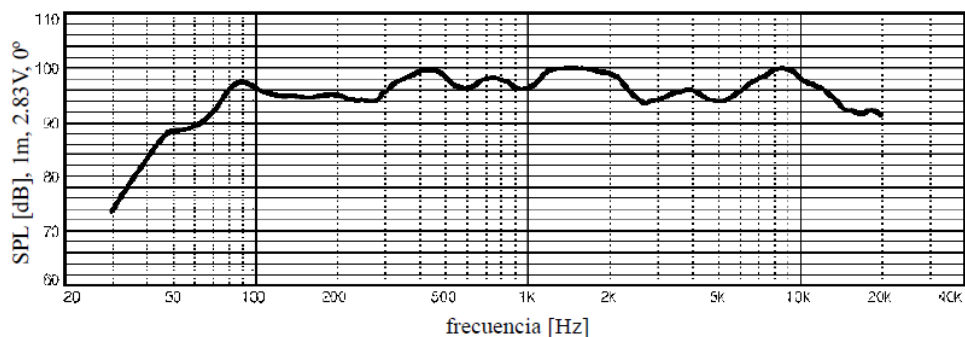


Figura 2-40. Respuesta en frecuencia obtenida de un barrido de tonos.

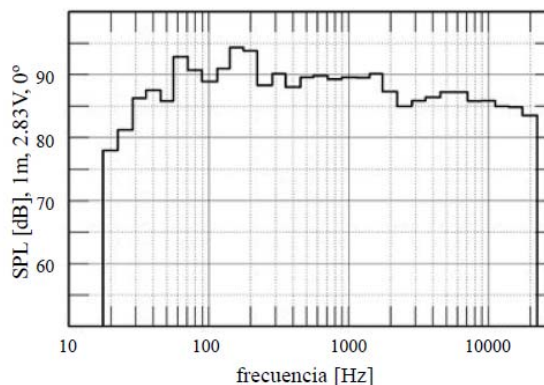


Figura 2-41. Respuesta en frecuencia obtenida de excitación con tercios de octava.

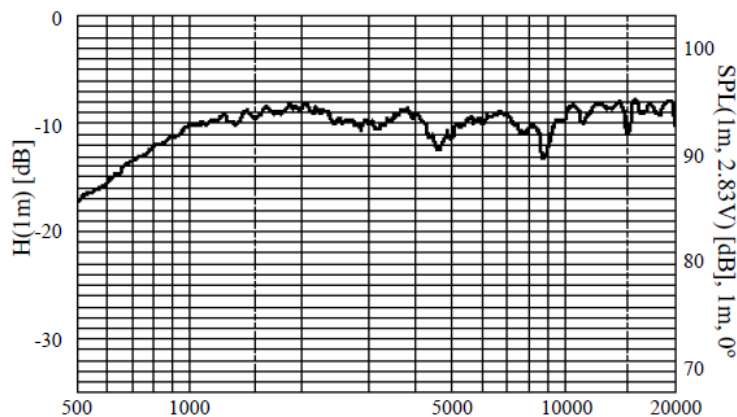


Figura 2-42. Respuesta en frecuencia obtenida de la función de transferencia del altavoz.

2.4.5 Directividad

Diagrama polar de directividad.

Refleja el nivel relativo a la radiación máxima, en función del ángulo de desviación del eje del altavoz. Se proporciona en forma de diagrama polar para cada frecuencia en el que se representa la directividad $D(\theta, \phi)$ [dB]. Se debe medir en cámara anecoica con bandas de ruido rosa, de un tercio de octava o una octava. La variable angular que mide el ángulo de elevación respecto del eje del altavoz. El ángulo de azimut (ϕ) usado normalmente es de 0° o 90° . En las Figura 2-43 se muestran la configuración geométrica para las curvas polares de directividad.

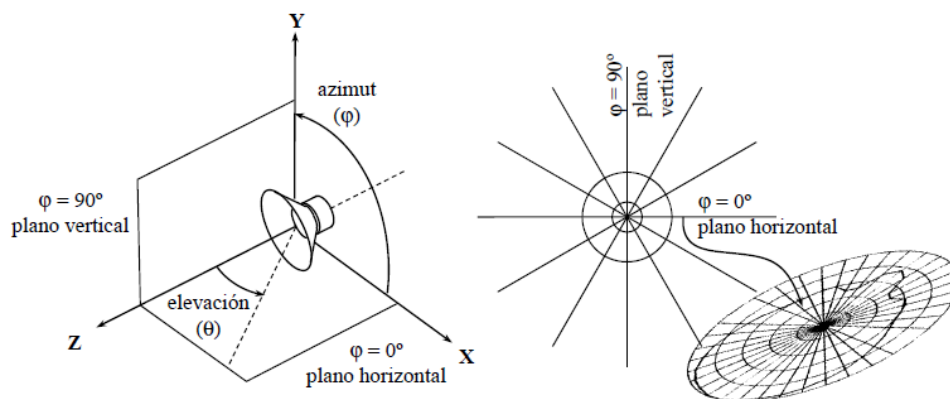


Figura 2-43. Configuración geométrica para las curvas polares de directividad. Planos horizontal y vertical y ángulos de elevación y azimut (izqda.). Juego de diagramas polares para diferentes ángulos de azimut recorriendo el intervalo ϕ desde 0° a 180° .

Por definición la directividad siempre está normalizada a 0dB (el valor máximo de las curvas siempre es el mismo), por lo que todos sus valores son negativos.

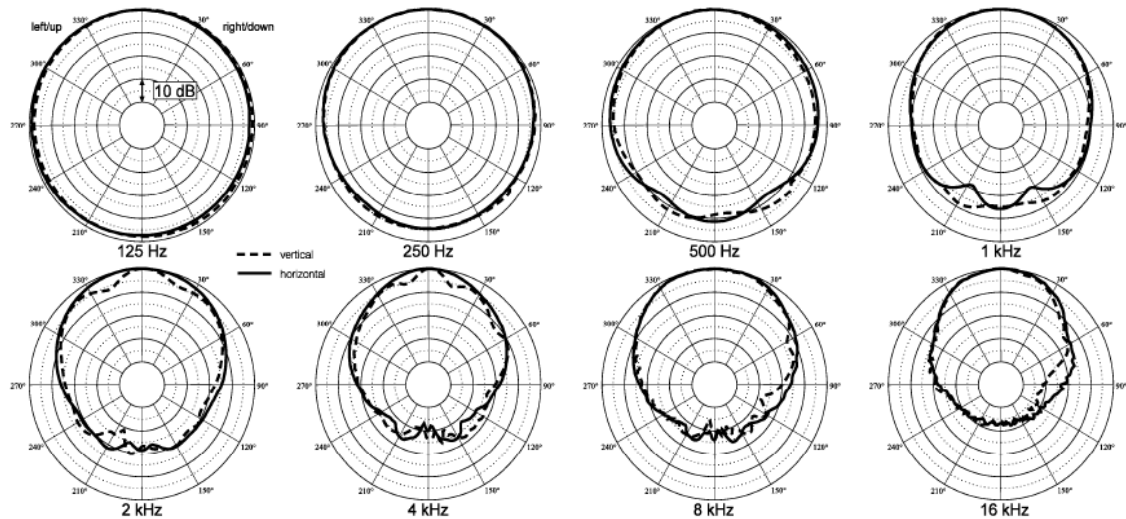


Figura 2-44. Diagramas polares de directividad medidos en 1/3 de octava, para una caja acústica de banda ancha. Para cada octava de frecuencia se indican las curvas de directividad horizontal y vertical de la caja.

Factor de directividad Q e índice de directividad DI

Aunque no es muy frecuente, algunos fabricantes especifican el factor de directividad axial Q_{ax} y el índice de directividad axial DI_{ax} , bien en forma de curva con la frecuencia o en forma de simples valores tabulados, normalmente en bandas de octava. El índice de directividad es la expresión logarítmica del factor de directividad.

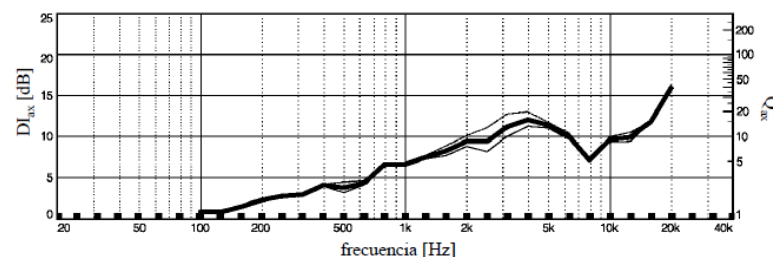


Figura 2-45. DI_{ax} y Q_{ax} en función de la frecuencia.

Ángulos de cobertura

Se especifica el ángulo en el que la pérdida por directividad es de 6 dB, tanto en el plano horizontal θ_{LH} (-6 dB) como en el vertical θ_{LV} (-6 dB). El ángulo de cobertura es una característica que se suele especificar especialmente en altavoces de bocina. Así una bocina de $60^\circ \times 40^\circ$, tiene un ángulo de cobertura horizontal de 60° y un ángulo de cobertura vertical de 40° .

Hay fabricantes que especifican además los ángulos de cobertura a -3 dB y a -9 dB, como el ejemplo de la Figura 2-46.

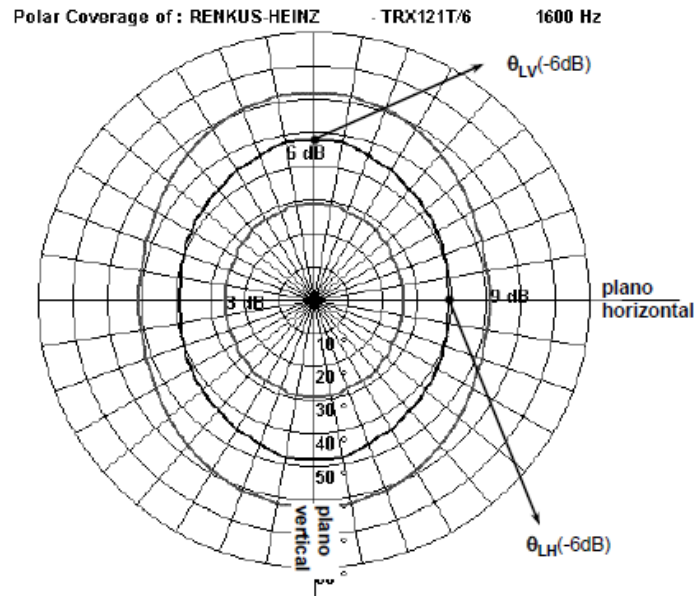


Figura 2-46. Ángulos de cobertura a -3, -6 y -9 dB para todos los planos del altavoz.

También es frecuente que se representen los ángulos de cobertura en función de la frecuencia, como se aprecia en el ejemplo de la Figura 2-47.

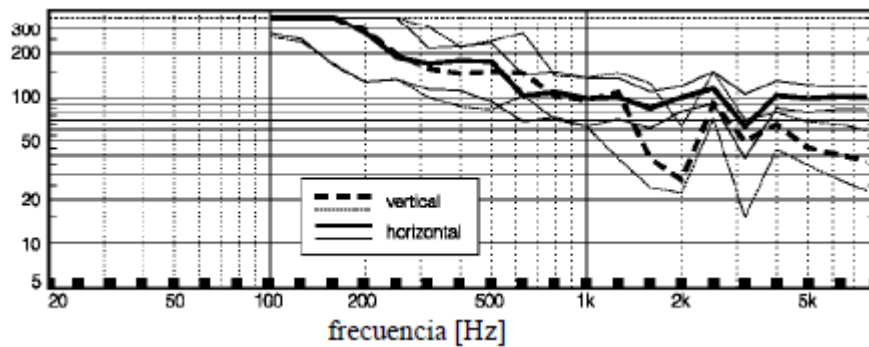


Figura 2-47. Ángulos de cobertura para -3, -6 y -9 dB. El trazo más grueso corresponde a los -6dB.

2.4.6 Eficiencia

El parámetro eficiencia, $\eta[\%]$ indica la relación entre la potencia acústica radiada por un altavoz, P_a , y la potencia eléctrica total consumida en los terminales de entrada, P_e , y se calcula con la expresión (2-6):

$$\eta[\%] = 100 \frac{P_a}{P_e} \quad (2-6)$$

Hay que recordar que la potencia eléctrica se calcula mediante (2-2). La eficiencia de un altavoz nunca supera el 50% y generalmente es menor al 10%. La potencia acústica P_a puede obtenerse en condiciones de campo libre (medida en cámara anecoica) o en condiciones de campo difuso (medida en cámara reverberante).

En condiciones de campo libre la expresión de cálculo P_a para radiación esférica es la (2-7) :

$$P_a = \frac{r^2}{\rho_0 c} \int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\pi} p^2(\theta, \varphi) \sin\theta \, d\theta \, d\varphi \quad P_a. (2-7)$$

Con r la distancia y $p(\theta, \varphi)$ la presión medida para diferentes ángulos de elevación y azimut. En condiciones de radiación hemisférica (cámara semi-anecoica) la ecuación (2-7) se transforma en (2-8):

$$P_a = \frac{r^2}{\rho_0 c} \int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\pi/2} p^2(\theta, \varphi) \sin\theta \, d\theta \, d\varphi \quad \text{Pa. (2-8)}$$

Cuando se conoce la sensibilidad del altavoz y el factor de directividad axial, Q_{ax} , la potencia acústica se puede calcular con las expresiones (2-9) y (2-10):

-Para radiación esférica:

$$P_a (P_e = 1W) \equiv \eta = 10^{\frac{S - 10 \cdot \log Q_{ax} - 109}{10}} \quad \text{Pa. (2-9)}$$

Donde $-109 = 10 \cdot \log \frac{4\pi p_{ref}^2}{\rho_0 c}$

-Para radiación hemisférica:

$$P_a (P_e = 1W) \equiv \eta = 10^{\frac{S - 10 \cdot \log Q_{ax} - 112}{10}} \quad \text{Pa. (2-10)}$$

Donde $-112 = 10 \cdot \log \frac{2\pi p_{ref}^2}{\rho_0 c}$

2.4.7 Distorsión

Ya se comentó en el apartado de características constructivas que la configuración, materiales y disposición de los elementos que conforman el altavoz tenían influencia en la distorsión del mismo. En la Figura 2-48 se muestra un ejemplo de distorsión en función de la frecuencia de un altavoz comercial.

La forma más habitual de evaluar la distorsión en un altavoz es excitarlo con un tono puro y observar el nivel de armónicos recogidos por un micrófono a cierta distancia. De esta manera se conoce la distorsión armónica (HD o *Harmonic Distortion*). La distorsión medida sería la relación entre la tensión RMS de los armónicos de distorsión respecto al armónico principal. Se puede medir la distorsión armónica total o la correspondiente a cada armónico (THD o TH2, TH3, etc). Para obtener el valor de distorsión armónica hay que excitar el altavoz con un tono puro a diferentes frecuencias, con un nivel de 1 W o un porcentaje de la potencia nominal del altavoz.

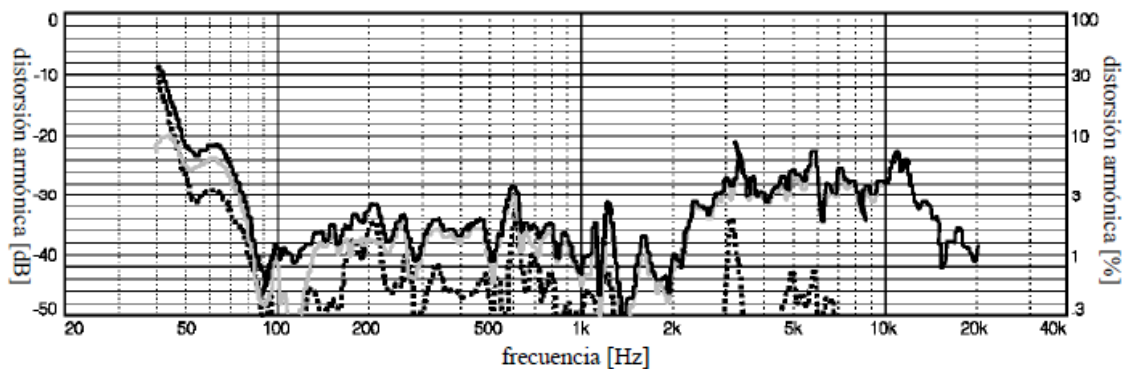


Figura 2-48. Distorsión armónica en función de la frecuencia de un altavoz comercial excitado con tonos puros del 10% de su potencia nominal senoidal. En línea continua THD+ruido, en gris TH2 y en línea discontinua TH3.

2.4.8 Parámetros de Thiele-Small.

Son los parámetros que determinan el modelo electromecánico de un altavoz dinámico. Su nombre procede de A. N. Thiele y R.H Small, dos ingenieros que de forma independiente propusieron dicho modelo alrededor de los años 70.

Los parámetros que se definen están medidos en pantalla infinita aunque también existen parámetros de Thiele-Small de caja cerrada.

2.5 Cajas acústicas

Si disponemos de un altavoz y lo alimentamos con una señal eléctrica, se generan ondas sonoras por las dos caras de su diafragma. Para evitar que estas señales se anulen, debemos situar los altavoces dentro de unas cajas, que serán de madera o plástico, y que llamaremos bafles o cajas acústicas.

- Además, debemos tener en cuenta las siguientes consideraciones:
- Los altavoces deben atornillarse firmemente a las cajas para que no provoquen vibraciones en éstas.
- La caja debe ser rígida.

Deben situarse a una altura de unos 40 ó 50 cm, lo que mejora la respuesta de los agudos.

Caja infinita o cerrada

Para conseguir que las ondas no se anulen se utiliza una caja a la que se le practica un agujero en el que se coloca el altavoz. Se caracterizan por la pérdida de graves aunque, en parte, podemos evitar este problema si recubrimos el interior de la caja con un material absorbente.

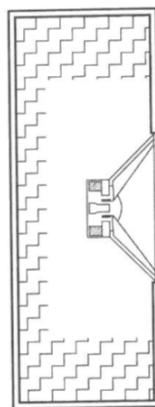


Figura 2-49. Caja acústica cerrada.

Caja Bass-reflex

Es muy similar a la anterior, pero se practican más agujeros que permiten recuperar parte de la energía acústica del interior.

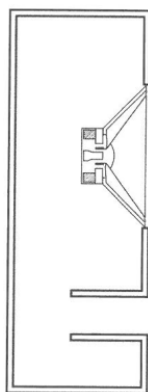


Figura 2-50. Caja acústica abierta con un sistema de Bass-reflex.

Refuerza la respuesta en bajas frecuencias. Si alguna de las ventanas adicionales está en la parte posterior de la caja, debemos separarla de la pared un mínimo de 10 cm. Además, la pared opuesta debe estar a una distancia mínima de 1 m. para que todas las frecuencias estén más equilibradas.

Caja con radiador pasivo o activo-pasivo

Es un caso particular del anterior, pero en esta ocasión se coloca en un nuevo hueco un altavoz sin alimentación eléctrica que llamamos radiador. Este elemento aprovecha las ondas que genera el diafragma por la parte posterior y las emite. Mejora la respuesta de graves y permite reducir la distorsión de la onda sonora.

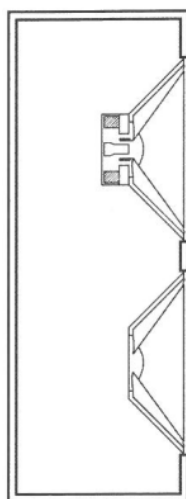


Figura 2-51. Caja con radiador pasivo.

2.5.1 Conexión de las cajas acústicas

Para conectar las cajas acústicas a un sistema de audio, puede emplearse hilo pelado o conectores tipo banana.

En cualquier caso, siempre existen dos colores; el rojo para el positivo y el negro o blanco para el negativo. Cuando conectemos las cajas al amplificador debemos tener en cuenta esto, porque una conexión incorrecta provoca una pérdida de graves.

En usos profesionales, se utilizan los conectores SpeakOn, que se conectan igual que los BNC, dando un cuarto de vuelta, aunque también incorporan una pestaña de seguridad como los XLR.

Cuando la distancia entre la fuente de la señal eléctrica y las cajas sea superior a 10 m, es recomendable utilizar un cable blindado de calidad para que las pérdidas sean lo más pequeñas posibles.

Monitores:

Son las cajas que se utilizan en los estudios de grabación. Podemos diferenciar dos tipos:

- **Monitores de alta potencia:** Se suelen colocar en las esquinas superiores de la pared más estrecha. Se ven afectados por la acústica de la sala y permiten escuchar, con mucha precisión, el resultado final de la mezcla.
- **Monitores de baja potencia:** Se sitúan al lado del técnico y permiten escuchar la mezcla tal y como se haría de forma doméstica.

Sistemas de escucha:

Con la aparición de los *home cinema*, se incorporan nuevos equipos que utilizan un mínimo de cinco altavoces.

Estos sistemas se caracterizan por:

- Disponer de un altavoz de canal central, que se sitúa encima o debajo del monitor de televisión.
- Incorporar altavoces de efecto (*surround*), que mejoran los efectos especiales.
- Disponer de pre-ajustes que permiten escuchar cualquier fuente de sonido en determinados ambientes (grandes teatros, salas pequeñas, etc.).

En la Figura 2-52, se nos muestra la disposición típica de uno de estos equipos.

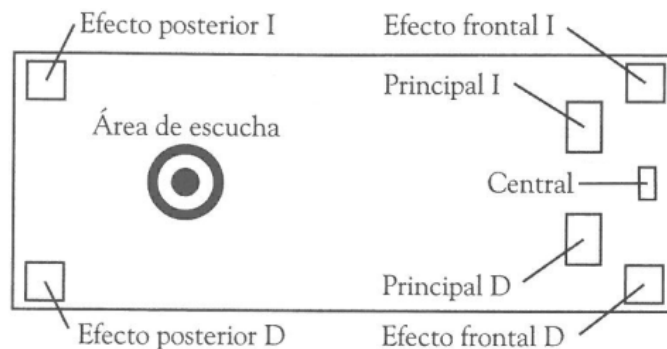


Figura 2-52. Sistema sonoro de un *home cinema*.

Cuando tenemos altavoces separados por una distancia considerable como, por ejemplo, en un concierto, debemos intentar que el sonido llegue correctamente a todas las ubicaciones del público. Para conseguir este efecto, se aplican **retardos** de la señal de audio a los distintos grupos de altavoces, siendo éste mayor a medida que nos alejamos de los altavoces frontales.

2.6 Cuestionario

1-Los altavoces con diafragma pequeño no pueden radiar en graves, pero un altavoz de gran diafragma si puede radiar en agudos.

- a) Verdadero.
- b) Falso.

2-La disipación de calor en el interior del altavoz es un factor que influye en su sensibilidad, potencia y eficiencia.

- a) Verdadero.
- b) Falso.

3-La sensibilidad, S , de un altavoz:

- a) es el nivel de presión sonora medido en el eje y a 1 metro de distancia, en cámara reverberante, cuando se le excita con ruido rosa filtrado en su banda útil y con una potencia de $P_e = 1W$. Se especifica en dB.
- b) Es el nivel de presión sonora medido en el eje y a 1 metro de distancia, en cámara anecoica, cuando se le excita con ruido rosa filtrado en su banda útil y con una potencia de $P_e = 1W$. Se especifica en dB.
- c) Está representada por su curva de nivel en función de la frecuencia y medida en campo libre (cámara anecoica), obtenida por barrido de tonos o en tercios de octava, todo a 1W.
- d) Ninguna de las anteriores es cierta.

4-Calcular la tensión a la que equivale una señal eléctrica de 1W desarrollada sobre 2 ohmios. Respuesta con solo 2 decimales.

5-Los preamplificadores sirven para variar la ganancia de los altavoces, evitando así que radien en exceso.

- a) Verdadero.
- b) Falso.

6- Los altavoces coaxiales y los de doble cono:

- a) Se diferencian en la transducción electroacústica. Mientras que los de doble cono solo tienen una bobina para sus dos diafragmas, los coaxiales constan de circuitos magnéticos independientes para cada diafragma, además de incluir un crossover para la división del espectro.
- b) Se diferencian en la transducción electroacústica. Mientras que los coaxiales solo tienen una bobina para sus dos diafragmas, los de doble cono constan de circuitos magnéticos independientes para cada diafragma, además de incluir un crossover para la división del espectro.
- c) Son denominaciones diferentes para designar a los altavoces elípticos, también denominados de banda ancha.
- d) Ninguna de las anteriores es cierta.

7- Los altavoces:

- a) Constan de dos transductores que transforman la señal eléctrica (señal de audio) en señal acústica.
- b) Constan de dos transductores, colocados en el mismo orden que en los micrófonos, que transforman la señal eléctrica en señal acústica.

c) No pueden incluir en el interior de su caja acústica filtros crossover ya que se producirían problemas eléctricos que afectarían en la respuesta en frecuencia del altavoz.

d) Ninguna de las anteriores es correcta.

8-La respuesta en frecuencia de un altavoz:

a) Es el nivel de presión sonora medido en el eje y a 1 metro de distancia, en cámara anecoica, cuando se le excita con ruido rosa filtrado en su banda útil y con una potencia de $P_e = 1W$, especificado en dB.

b) Se obtiene a partir de la función de transferencia del altavoz, $H(r)$, para evitar los problemas de niveles bajos obtenidos con la obtención a partir de ruido en banda ancha.

c) Refleja el nivel relativo a la radiación máxima, en función del ángulo de desviación del eje del altavoz.

d) Ninguna de las anteriores es correcta.

9- La rigidez y la ligereza del material con el que se fabrican los diafragmas:

a) Afecta de tal manera a la respuesta del altavoz, que se han de buscar materiales rígidos y ligeros para aumentar la eficiencia y sensibilidad del altavoz.

b) Tienen un peso importante en la respuesta del altavoz, pero es la manera de colocación (plana, con corrugaciones o curva) la que afecta de forma decisiva en la sensibilidad del altavoz.

c) No tienen importancia relevante en la respuesta del altavoz. Es más importante el material con el que se fabrican las suspensiones sobre las que se soporta el diafragma.

d) Ninguna de las anteriores es correcta.

10- Debido a las características de los elementos que conforman el altavoz

a) Es muy simple diseñar sistemas que permitan radiar en todo el margen de audio con un solo altavoz.

b) No es posible diseñarlos con un diámetro mayor de 12 pulgadas ya que generarían modos propios transversales y circunferenciales a cualquier frecuencia.

c) Es necesario el diseño de filtros de cruce, para dividir el espectro y enviar a cada altavoz la banda que mejor radie.

d) Ninguna de las anteriores es correcta.

11-Las cajas acústicas en las que se incrustan los altavoces sirven únicamente como soporte para éste y no aportan ningún matiz al sonido final.

a) Verdadero.

b) Falso.

12- Los parámetros de Thiele-Small:

a) Son los parámetros que determinan el modelo electromecánico de un altavoz dinámico.

b) Son los parámetros que determinan el modelo electromecánico de un altavoz coaxial.

c) Son los parámetros correspondientes al circuito eléctrico del filtro pasivo de un altavoz de dos vías.

d) Ninguna de las anteriores es cierta.

13- En la distorsión que se produce en los altavoces:

a) No influyen las características constructivas.

b) Influyen las características constructivas de la configuración, materiales y disposición de los elementos que conforman el altavoz.

- c) Entren en juego los parámetros de Thiele-Small, al controlar sus valores se puede controlar la distorsión.
- d) Ninguna de las anteriores es cierta.
- 14- Los altavoces cuentan con dos suspensiones, la primaria, elemento responsable de dar elasticidad al movimiento del diafragma, y la secundaria, encargada de sujetar la parte superior del diafragma a la estructura del altavoz.
- a) Verdadero.
- b) Falso.
- 15- La bobina del altavoz se une al *former* con cemento termo-resistente.
- a) Verdadero.
- b) Falso.
- 16- Los altavoces se comportan a la vista de un amplificador de potencia como:
- a) Un condensador.
- b) Un transistor.
- c) Una impedancia de línea de 600 ohmios.
- d) Ninguna de las anteriores es cierta.
- 17- El filtrado:
- a) Activo se implementa de tal manera que cada zona del espectro tiene la misma amplificación.
- b) Pasivo se implementa de tal manera que los filtros van incluidos en las cajas acústicas con el altavoz y se alimentan con un único amplificador de potencia.
- c) Activo y pasivo se diferencia en la estructura interna de los filtros, las respuestas en frecuencia es muy similar pero la salida del filtrado activo, a diferencia del pasivo, genera armónicos en las componentes pares de la señal.
- d) Ninguna de las anteriores es cierta.
- 18- Una mesa de mezclas es un complejo sistema de audio, que no sólo preamplifica y asigna señales de entrada a diferentes salidas, también permite una variedad de diferentes mezclas, un enrutamiento y procesamiento especial de la señal de audio.
- a) Verdadero.
- b) Falso.
- 19- Los altavoces dinámicos de cinta se basan en el mecanismo del condensador variable, usado también en los micrófonos de condensador. Su estructura consiste en tres placas metálicas que forman una especie de bocadillo. Las dos placas exteriores, rígidas y perforadas, constituyen el condensador. La placa interna, mucho más delgada, es el diafragma.
- a) Verdadero.
- b) Falso.
- 20- Los altavoces de graves:
- a) Son los que normalmente reproducen frecuencias bajas, hasta los 500 Hz, dependiendo del caso.
- b) Están especializados en radiar frecuencias por encima de unos 5 kHz.

- c) Reproducen la banda media de frecuencias, desde los 500 hasta los 5kHz.
- d) Ninguna de las anteriores es cierta.

3 CÁMARAS Y SENSORES

Las cámaras de vídeo son equipos que captan una imagen a partir de la realidad circundante y la convierten en señal eléctrica. Es decir, una cámara de vídeo es un transductor óptico-eléctrico. La cámara usa unos sensores sensibles a la luz que incide a través del objetivo, de forma que mediante un muestreo o escaneo concreto (dictaminado por normativas), se obtienen señales eléctricas unidimensionales y continuas de tensión, $V(t)$.

3.1 Introducción a la señal de vídeo

Para afirmar que se está viendo algo debe haber luz, es decir, en la escena donde nos encontremos debe haber algún emisor de ondas electromagnéticas emitiendo dentro del espectro visible. En la Figura 3-1 se representa de forma esquemática el espectro electromagnético conocido.

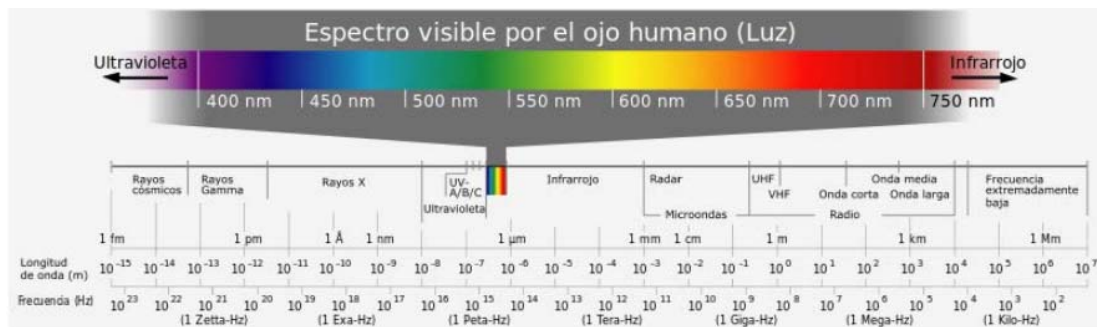


Figura 3-1. Espectro electromagnético conocido.

El espectro visible es una pequeña zona en el espectro electromagnético, que va desde los 750 nm a los 400 nm.

El máximo de respuesta del ojo se encuentra en la longitud de onda de 555 nm que corresponde a una radiación que provoca la sensación de verde. Para valores mayores o menores a ese la respuesta del ojo decrece. En la Figura 3-2 se muestran las curvas de respuesta del ojo. La diferencia entre las dos curvas se debe a que el ojo tiene distintos tipos de sensores que se excitan con intensidades de luz medias o altas, dando origen a respuestas algo diferentes. En entornos de vídeo se trabaja con iluminaciones medias o altas.

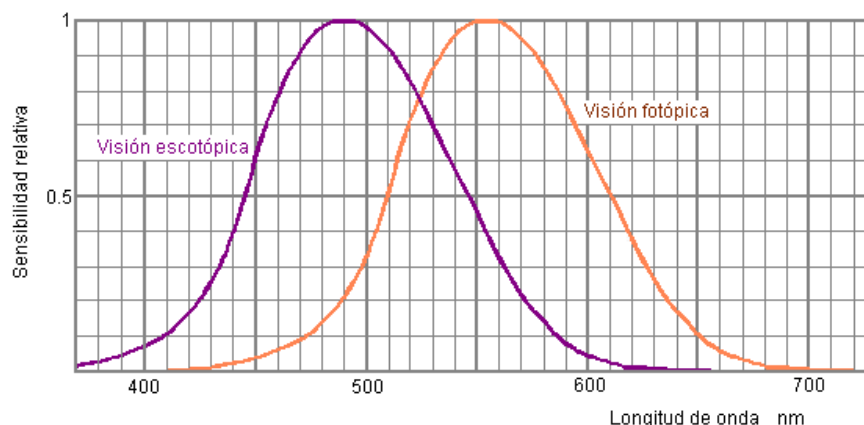


Figura 3-2. Curvas de respuesta del ojo. Visión escotópica = niveles bajos de iluminación y visión fotópica = niveles altos de iluminación.

Cuando una luz llega al ojo, incidiendo en sus sensores, se produce sobre el cerebro una sensación de luz coloreada que tiene tres características:

Brillo: Es la cantidad de luz que percibe el ojo y tiene una relación directa con la luminancia. El contraste se define como la diferencia entre el nivel máximo y mínimo de luz que existe en una imagen.



Figura 3-3. Niveles de brillo. De mayor a menor de izquierda a derecha.

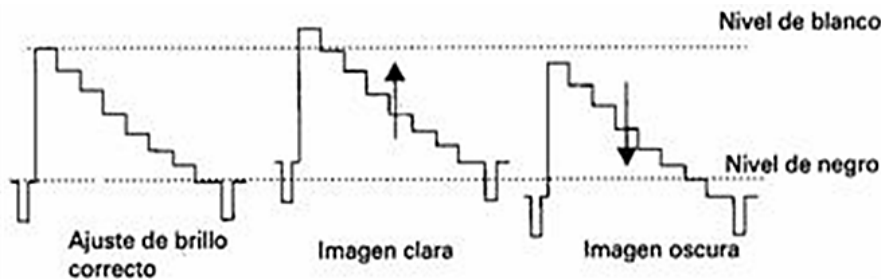


Figura 3-4. Efecto de la variación en el brillo.

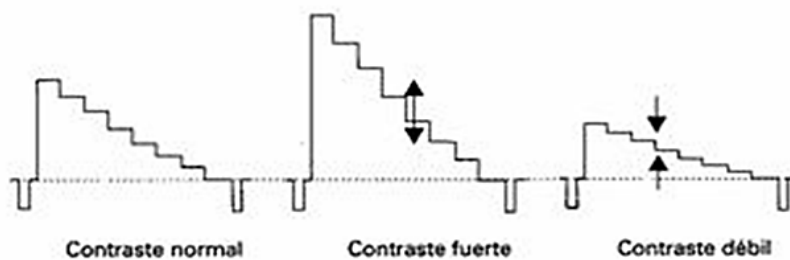


Figura 3-5. Efecto de la variación en el contraste.

Matiz o tono: Es la propiedad de la luz que permite determinar el color de un objeto y está relacionado con la longitud de onda dominante.

Saturación: Está relacionado con el nivel de color o cantidad de color que produce una imagen.

3.1.1 Mezcla aditiva espacial. Agudeza visual

El número de sensores colocados sobre la retina es muy elevado, en torno a 130 millones. Se diferencian en que unos se excitan cuando hay muy poca iluminación, cilindros, y otros se excitan cuando hay mucha, conos. Los cilindros (unos 120 millones) se sitúan en la parte externa de la retina, mientras que los conos (unos 8 millones) se sitúan en la zona central. Aunque el número es elevadísimo, no tenemos un número infinito de sensores, lo cual produce un efecto denominado mezcla aditiva espacial, que a continuación se explica.

Si dos puntos de una imagen se encuentran lo bastante separados afectarán a dos sensores diferentes y, por tanto, serán dos informaciones diferentes las que procesará nuestro cerebro, como se aprecia en el ejemplo de la Figura 3-6. En cambio, si los dos puntos están muy juntos, cosa que ocurre si nos alejamos de ellos, puede que solo sean captados por un sensor y, por tanto, la percepción será como si viésemos sólo un punto y no dos diferenciados, como se observa en la Figura 3-7. Cuando esto sucede se dice que se ha efectuado la mezcla aditiva espacial, ya que aunque dos puntos estén separados, nuestro sistema visual los mezcla y los percibimos como uno solo.

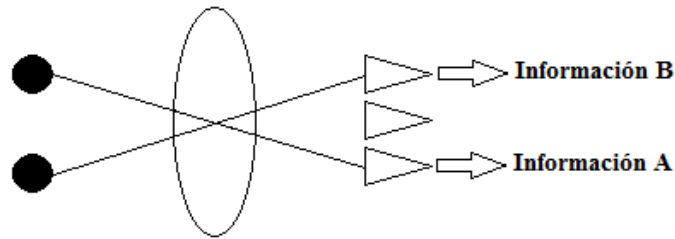


Figura 3-6. Captación de dos puntos diferenciados.

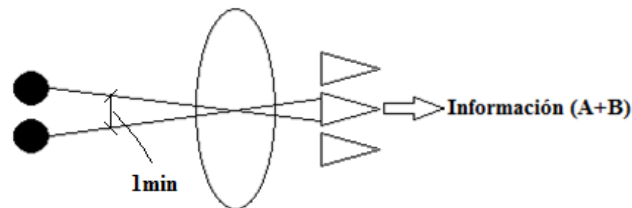


Figura 3-7. Captación de dos puntos diferenciados como uno solo. Límite entre dos puntos respecto al ojo para que esto ocurra: 1' de grado.

Se puede enfocar este tema diciendo que la **agudeza visual** de ojo no es infinita, sino que se puede cuantificar. Y realizando experimentos se ha llegado a comprobar que para un ojo de características medias, la agudeza visual es de un minuto de grado, si dos puntos se ven con un ángulo menor se percibirán como uno solo. Esta limitación del ojo se aprovecha también para conseguir que el ojo mezcle colores de distintos puntos. De esta manera, con la mezcla de los colores primarios (rojo, verde y azul (RGB)) se puede hacer que percibamos infinidad de tonalidades cromáticas. 5 En la televisión en color, cada punto de una imagen o píxel está formado por 3 luminóforos, uno rojo, otro verde y otro azul, como se puede ver en la Figura 3-8. Jugando con la luminosidad de cada uno de ellos se pueden obtener multitud de colores sin que percibamos la imagen como una suma de puntos, sino como una imagen continua en toda su superficie.

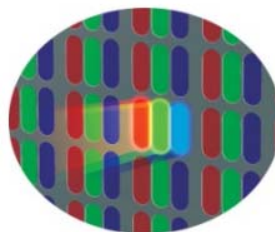


Figura 3-8. Posiciones relativas de los luminóforos en las pantallas de televisión.

3.1.2 Mezcla aditiva temporal. Memoria visual

El ojo es muy rápido en su respuesta y puede pasar de un estado de excitación a otro de reposo casi instantáneamente. Pero el tiempo que tarda en responder al cambio no es cero. Si se excita al ojo con un impulso luminoso, cuando el impulso ha desaparecido los sensores del ojo siguen mandando información al cerebro, lo cual se traduce en que se sigue teniendo sensación luminosa, como se puede apreciar en el ejemplo de la Figura 3-9.

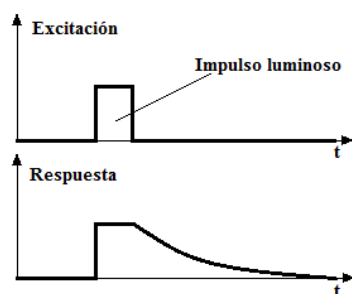


Figura 3-9. Respuesta del ojo a un impulso.

Si antes de que la respuesta al impulso desaparezca totalmente se excita el ojo con otro impulso, no habrá sensación de que ha habido dos impulsos sino solo uno.

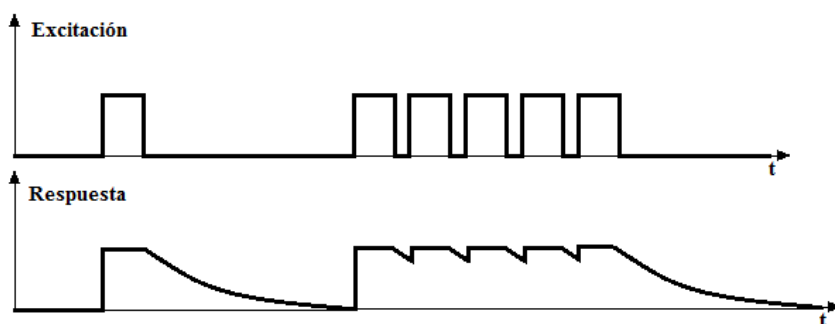


Figura 3-10. Respuesta del ojo a varios impulsos seguidos.

Este fenómeno se conoce como mezcla aditiva temporal, ya que dos impulsos que se producen en diferentes instantes de tiempo son integrados por el ojo y los considera unidos en el tiempo, sin discontinuidad entre ellos.

Si la distancia temporal entre impulsos es intermedia, no se aprecia una separación total entre ellos pero sí un efecto conocido como parpadeo o *flicker*.

El tiempo dentro del cual el ojo integra dos impulsos como uno es de 50 ms, valor que se ha usado desde que se inventó el cine. Esto se traduce en que para provocar la sensación de movimiento continuo, sin parones, es necesario usar un número determinado de impulsos por segundo, o lo que es lo mismo un número determinado de imágenes por segundo. Inicialmente se usaron 24 fotogramas por segundo (42 ms por fotograma) en cine. Si el ojo no tuviese esta memoria visual sería necesario proyectar infinitas imágenes por segundo para crear la sensación de movimiento, y aun así se percibiría parpadeo entre imagen e imagen. En los sistemas americanos se optó por usar 30 imágenes por segundo (33 ms por imagen) y en sistemas europeos se decantaron por 25 imágenes por segundo (20 ms por imagen). La elección del número de imágenes por segundo se tomó por aspectos técnicos relacionados con la frecuencia de la red eléctrica.

3.2 Sistema de transmisión de imágenes

Una vez conocido el número de imágenes necesarias para crear la sensación de movimiento y la agudeza visual del ser humano, hay que saber cómo convertir esa sucesión de imágenes en una señal eléctrica. Es decir, hay que saber cómo transformar un parámetro bidimensional, la imagen, en un señal unidimensional, tensión. En 1881 un notario francés, Constantino Senlecq, tuvo la idea de aprovechar la memoria visual para crear un sistema secuencial de transmisión, base de todos los sistemas actuales. Esta idea se basa en que el ojo no necesita tener constantemente información de todos los puntos de la escena, basta con que dicha información se le vaya suministrando cada cierto tiempo, pero siempre dentro de esos 50 ms que se comentaron anteriormente. De esta manera el ojo no notará discontinuidad en la escena. Aunque es necesario una elevada velocidad a la hora de excitar/no excitar es posible implementar este tipo de sistemas. Se denomina sistema secuencial porque se transmite secuencialmente la información de cada uno de los puntos, uno detrás de otro, hasta acabar transmitiéndolos todos. Esto simplifica enormemente los circuitos existentes entre cámara y monitor, ya que se

puede trabajar con una sola línea de transmisión. El problema que se encontraron al llevarlo a la práctica fue la rapidez de exploración con la que se tenía que trabajar, pero actualmente con los dispositivos electrónicos de los que se dispone esto no supone ningún obstáculo.

3.2.1 Número de líneas

Se ha comentado que por la agudeza visual humana, dos puntos que formen un ángulo inferior a 1 minuto de grado con respecto al ojo serán percibidos como uno solo (para las señales de color este ángulo es de unas tres a cinco veces mayor). Lo que da pie a plantearse el número de líneas necesarias en los sistemas de vídeo para que el espectador no aprecie la distancia entre las mismas.

Para establecer el número de líneas, hay que tener en cuenta dos aspectos contrapuestos: a mayor número de líneas, menos sensible será el ojo a su separación, y a mayor información captada, mayor es el ancho de banda necesario para transmitirla. La definición se define como el máximo número de líneas que se puede discernir sobre una pantalla, a una distancia igual a la altura de ésta, si bien se considera como distancia mínima de observación aquella igual a cuatro veces la altura de la pantalla. En la Figura 3-11 se especifica las medidas para el cálculo del número de líneas.

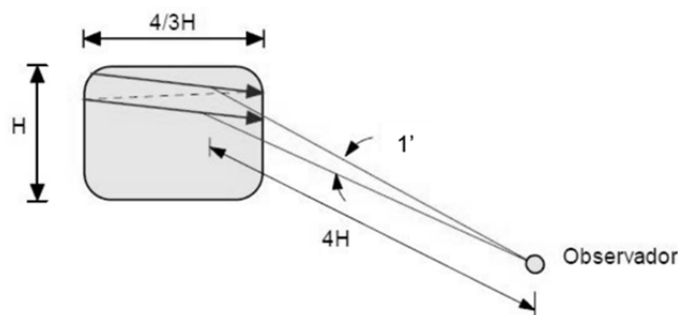


Figura 3-11 Cálculo del número de líneas.

La distancia entre líneas d se calcula con la expresión (3-1):

$$\tan(1') = \frac{d/2}{4H} = \frac{d}{8H}. \quad (3-1)$$

Luego el número de líneas mínimo que se establece para una imagen se calcula con la expresión (3-2):

$$N^{\circ} \text{ líneas} = \frac{H}{d} = \frac{1}{8 \cdot \tan(1')} = 430 \text{ líneas}. \quad (3-2)$$

En Europa el valor que se decidió tomar fue 625 superando el mínimo necesario para dejar algunas para el barrido vertical (como se verá más adelante) y siendo este valor múltiplo de 25 (imágenes por segundo o frecuencia de cuadro).

Dado que se trabaja con 25 imágenes por segundo con 625 líneas cada imagen, la **frecuencia de línea** se calcula con la expresión (3-3):

$$f_{\text{línea}} = Nv \cdot f_{\text{cuadro}} = 625 \cdot 25 = 15625 \text{ Hz}. \quad (3-3)$$

luego el tiempo asignado por el oscilador a cada una de las líneas viene dado por la expresión (3-4):

$$T_H = 1/f_{\text{línea}} = 1/15625 = 64 \mu\text{s}. \quad (3-4)$$

3.2.2 Ancho de banda

Para calcular el ancho de banda de la señal de vídeo hay que tener en cuenta varios parámetros. Uno de ellos es la relación de aspecto que es la relación entre la anchura y la altura de la imagen. La relación de aspecto que se usó originalmente fue 4:3. Para calcular el ancho de banda que ocupa una señal de vídeo nos ponemos en el peor

caso, es decir, imaginamos que tenemos que transmitir la máxima información posible. La imagen más compleja que se puede transmitir es una similar a un tablero de ajedrez como la que se aprecia en la Figura 3-12.

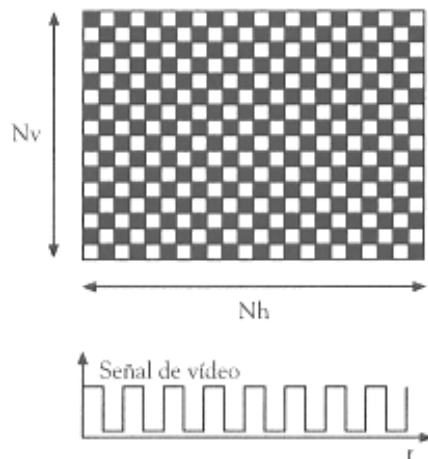


Figura 3-12. Señal transmitida para calcular la máxima frecuencia de variación (ancho de banda) de la señal de vídeo.

Se puede calcular el ancho de banda con la expresión (3-5) :

$$BW_{\text{vídeo}} = \frac{1}{T} = \frac{N_h/2}{T_H} = \frac{Z \cdot N_v/2}{1/f_{\text{línea}}} = \frac{Z \cdot N_v}{2} = \frac{N_v^2 \cdot Z \cdot f_{\text{cuadro}}}{2} \text{ Hz. (3-5)}$$

Siendo \$N_v\$ el número de líneas, \$Z\$ la relación de aspecto (\$N_h/N_v\$) y \$f_{\text{cuadro}}\$ el número de imágenes por segundo. Para \$N_v = 625\$, \$Z = 1,33\$ (4:3) y \$f_{\text{cuadro}} = 25\$ imágenes/s se obtiene un valor aproximado a 6,5 MHz. Para televisión analógica se estableció un valor de calidad-ancho de banda de 6 MHz.

3.2.3 Resolución

En sistemas de vídeo la resolución se define como la cantidad de líneas verticales que se pueden representar sobre la pantalla, en una distancia igual a la altura de la misma. Se mide en líneas y es un número entero. No se suele especificar si es resolución horizontal o vertical porque siempre se considera la resolución horizontal ya que es la única susceptible de ser modificada cambiando el ancho de banda de transmisión. La resolución vertical es independiente del canal de transmisión y viene dada por el número de líneas. El factor resolución se define como el cociente entre la resolución del sistema y el ancho de banda usado en su transmisión y se calcula con la ecuación (3-6).

$$\text{Factor de resolución} = \frac{\text{Resolución del sistema (en líneas)}}{\text{Ancho de banda utilizado (en Hz)}} \quad (3-6)$$

3.2.4 Sistema de exploración

Las cámaras o sensores pueden captar solo la luminancia de las imágenes, sin el color, o captar la información de crominancia R, G y B para cada uno de los píxeles. La información generada en cada píxel (elemento de imagen - *picture element*), es la que hay que enviar al receptor, y para ello habrá que extraerla del plano fotosensible donde se proyecta la escena. La extracción se hará de forma ordenada, de izquierda a derecha, formando una estructura de líneas horizontales, extrayendo las cargas línea a línea hasta acabar con la última de ellas, como se muestra en la Figura 3-13. Una vez extraída una imagen se procederá a las siguientes, para crear la sensación de movimiento.

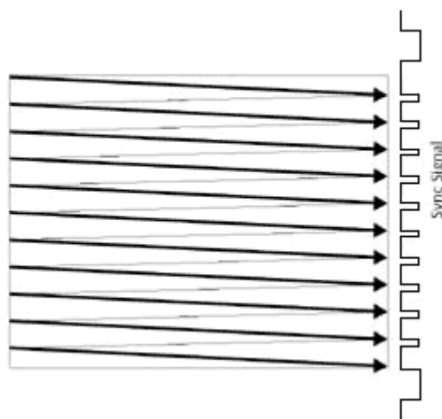


Figura 3-13. Sistema de exploración progresivo.

Cuando se acaba de analizar cada línea y cada imagen se plantea el problema de que, en la cámara y en el monitor se estén analizando y dibujando respectivamente puntos homólogos. Para salvaguardar este tema es necesario crear un sistema de sincronismos.

Al acabar de analizar una línea y mientras se está preparando para analizar la siguiente, se genera una señal denominada impulso de sincronismo horizontal (ISH), con características muy diferentes a la señal de vídeo para que sea fácilmente distinguible por el receptor. Con el ISH el receptor se adapta a la velocidad de llegada de cada línea. Al acabar la última línea de la imagen, debe avisarse al receptor que precisamente esa es la última línea. Este aviso se hace mediante el impulso de sincronismo vertical (ISV). Su misión es forzar al monitor a que la siguiente línea que dibuje sea la línea 1. A la señal de vídeo más los sincronismos se les denomina normalmente como **señal de vídeo**.

Frecuencia de cuadro

En vídeo, se denomina cuadro (*frame* en inglés) a una imagen completa. Frecuencia de cuadro es el número de imágenes transmitidas por segundo. En cine son 24 ips (imágenes por segundo). Lo lógico hubiera sido usar este mismo valor para la televisión, pero no fue así. La principal razón para que se eligiesen otros valores fue la frecuencia de red eléctrica, ya que es muy estable y si no se usase dicha frecuencia en los sistemas de vídeo aparecería ruido sobre la imagen. En Europa la frecuencia de red es de 50 Hz y el número de imágenes por segundo es 25, mientras que en América se eligió el valor de 30 imágenes por segundo, ya que su frecuencia de red es de 60 Hz.

Para una determinada frecuencia de cuadro, si el brillo de la imagen es pequeño, la sensación de parpadeo (*flicker*) es menor que si la imagen tiene un brillo considerable. Para determinar este efecto se desarrolló la denominada **ley de Ferry-Porter** que viene dada por la expresión (3-7):

$$f_c = 37 + 12,6 \cdot \log(B) \text{ Hz. (3-7)}$$

Para un determinado valor de brillo (B), la frecuencia de cuadro crítica por debajo de la cual se aprecia parpadeo es f_c .

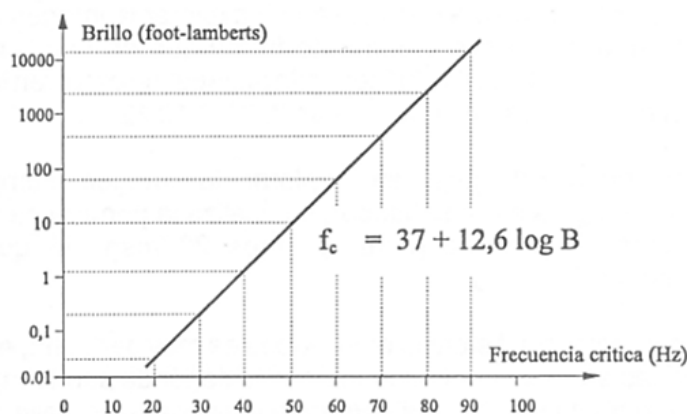


Figura 3-14. Ley de Ferry Porter.

Si la frecuencia de repetición de imágenes se establece en 25 imágenes por segundo, el brillo máximo que se le puede proporcionar a las pantallas sin que se aprecie parpadeo es de 0,11 foot-lamberts (unidad de la luminancia), que se puede afirmar que es un valor muy pequeño, de forma que si no se tomara ninguna medida, habría que mantener cerradas las ventanas para poder apreciar la imagen del televisor dentro de una habitación. Si se le diera más brillo a la pantalla, en las partes de la escena donde se superase ese valor de brillo, se produciría parpadeo en la imagen. Al duplicar la frecuencia de imagen, pasando a 50, el brillo que se podría forzar en la imagen sin apreciar parpadeo, pasa a ser de 10,8 foot-lamberts, donde puede comprobarse que ha pasado a ser 100 veces mayor que en la situación de 25 imágenes por segundo. Si se pasase a 100 imágenes por segundo, el brillo permitido sería de 100000 foot-lamberts o lo que es lo mismo, con esa frecuencia de imágenes no puede haber parpadeo.

Exploración entrelazada

Con 25 imágenes por segundo, el brillo máximo en la pantalla para que no se produzca parpadeo, según la ley de Ferry-Porter, es muy pequeño. Si se aumentase el brillo de la imagen se produciría dicho efecto no deseado. La solución más lógica sería aumentar el número de imágenes por segundo, pero esto supondría un aumento del ancho de banda de la señal de vídeo. La solución está en el método de explorar las imágenes. Si se exploran primero las líneas impares y después las pares se tienen dos semi-imágenes (campos - *fields*) que, para el ojo serán imágenes completas. Con esto se tienen dos imágenes donde antes solo había una, cada 40 ms se exploraría completamente la imagen y el ojo obtendría el doble de información: 50 imágenes por segundo aunque con menor resolución cada una de ellas, por lo que no sería necesario aumentar el ancho de banda. Este método de exploración es conocido como exploración entrelazada. En la Figura 3-15 se ejemplifica el funcionamiento del escaneo entrelazado de imágenes.

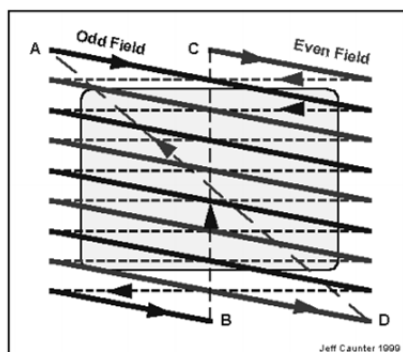


Figura 3-15. Escaneo entrelazado.

3.2.5 Corrección gamma

Presente en los sistemas de vídeo debido al comportamiento no lineal de alguno de los elementos que componen la cadena. En los receptores de televisión, cuando se utilizan tubos de rayos catódicos, no existe una relación del todo lineal entre la tensión que se aplica y el brillo que se genera en la pantalla, sino que es exponencial de la forma:

$$B_{salida} = V_{aplicada}^{\gamma}$$

Teniendo en cuenta esto, las emisoras de televisión corrigen en gamma las señales RGB antes de su emisión, de forma que de la suma de los dos efectos, la corrección y el efecto de los receptores, se obtenga una relación lineal.

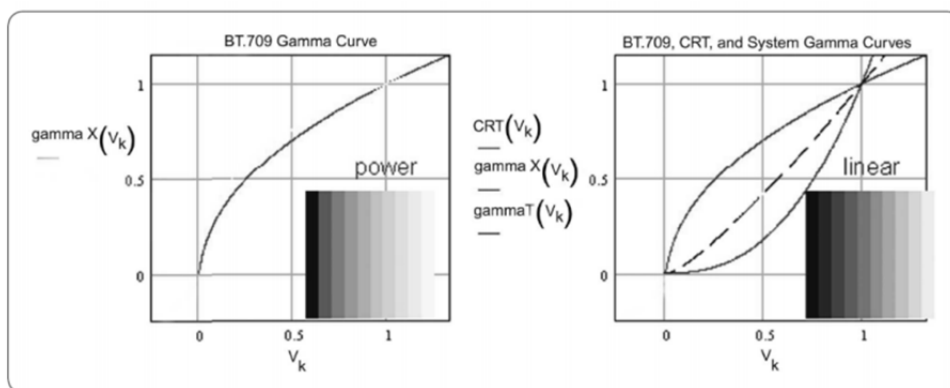


Figura 3-16. Curvas de corrección Gamma.

3.2.6 Sistemas de vídeo en color

Para reproducir una imagen o serie de imágenes en color es necesario que la cámara tenga sensores capaces de captarlos y los monitores tengan luminóforos capaces de reproducirlos. En la Figura 3-17 se muestra esquemáticamente la filosofía de funcionamiento de esta tecnología. A diferencia del sistema monocromo, que solo contaba con una señal por punto, en imágenes en color es necesario transmitir 3 señales por punto, que llevarán la información de la cantidad de rojo, azul y verde del punto en cuestión.

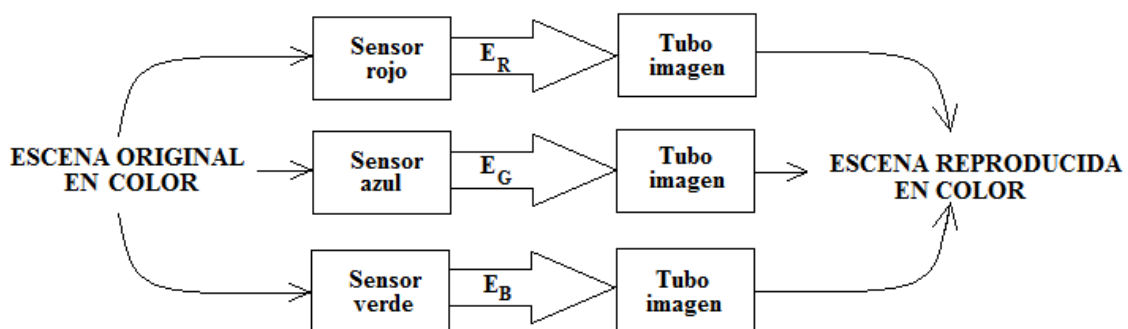


Figura 3-17. Esquema de la televisión en color.

Esto está basado en la Leyes de Grassman que dicen que por síntesis aditiva del color es posible conseguir todos los colores percibidos mezclando tres franjas del espectro visible (roja, verde y azul) en la proporción de intensidad adecuada. Siempre es posible variar la luminancia, manteniendo constante el matiz (el tono) y la saturación de color (la intensidad o mezcla con el blanco o negro).

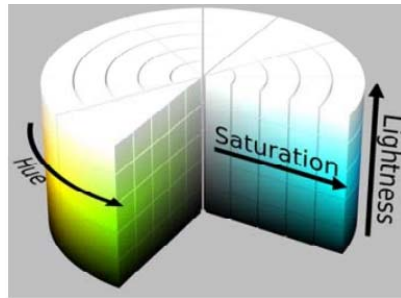


Figura 3-18. Representación de la luminancia, tono y saturación.

En 1931 la *Commission Internationale de L'Eclairage* (CIE) estableció tres señales, XYZ tristimulus, a partir de las cuales se pueden representar todos los colores por adición de ellas. En la Figura 3-19 aparecen representadas las funciones de dichas señales.

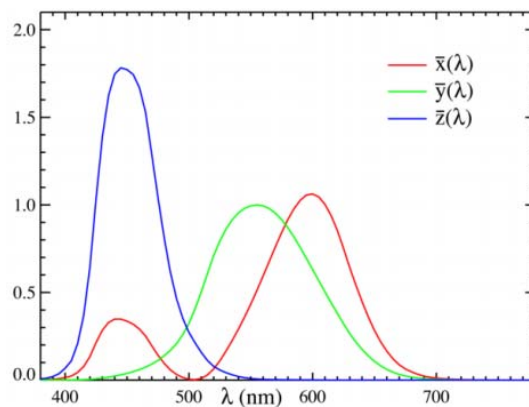


Figura 3-19. Funciones de XYZ CIE 1931.

Estas señales se normalizaron para poder representar todos los colores en dos dimensiones a través de las ecuaciones (3-8), (3-9) y (3-10):

$$x = \frac{X}{X+Y+Z}, (3-8) \quad y = \frac{Y}{X+Y+Z}, (3-9) \quad z = \frac{Z}{X+Y+Z}. (3-10)$$

Lo que llevó a la representación del diagrama de cromaticidad con todos los colores posibles, en la Figura 3-20.

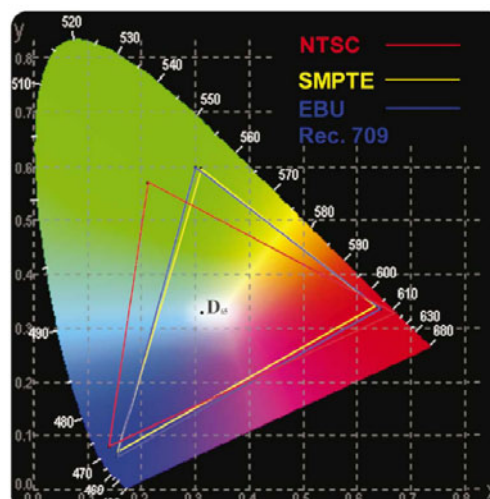


Figura 3-20. Diagrama de cromaticidad con los espacios de color de NTSC, SMPTE y EBU/PAL⁴.

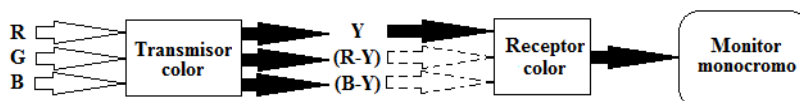
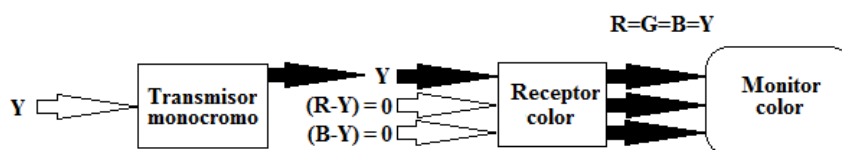
Los vértices de los triángulos representan los valores de RGB elegidos en cada espacio de color. Todos los colores incluidos son todos los posibles colores representables por ese espacio de color.

Se cuenta entonces con tres tensiones diferentes RGB, para lo cual sería necesario usar tres canales de transmisión diferentes. Pero debido a la circuitería y la necesidad de mantener la compatibilidad de los sistemas de color con los sistemas monocromo se optó por usar otro método de transmisión de la señal de vídeo en color, que consistía en enviar la luminancia (brillo) de cada punto, dada por la ecuación (3-11):

$$Y = 0,299R + 0,587G + 0,114B. (3-11)$$

De esta manera se respetaría la compatibilidad directa (que los sistemas monocromo puedan reproducir señales en color), representada en la Figura 3-21, y la compatibilidad inversa (que los sistemas de color puedan reproducir la señal monocroma), representada en la 3-22. Dado que de las cámaras recibimos todos los datos de la ecuación de luminancia, para transmitir la señal de color se optó por usar el método denominado de diferencias de color. Por este método se envía por un lado la luminancia, Y, y por otro las señales (R-Y) y (B-Y). La componente de color verde que falta por transmitir puede ser calculada por los decodificadores por la ecuación (3-12):

$$G = \frac{Y - 0,30R - 0,11B}{0,59}. (3-12)$$

**Figura 3-21. Compatibilidad directa.****3-22. Compatibilidad inversa.**

Luego en los sistemas de televisión no se suele transmitir RGB y se utiliza la señal en componentes Y, Pb y Pr calculadas partiendo de la ecuación (3-12), se obtienen las expresiones (3-13) y (3-14).

$$Pb = 0,564(B - Y). (3-13)$$

$$Pr = 0,713(R - Y). (3-14)$$

En la Figura 3-23 se puede ver las señales RGB que tienen un margen dinámico de 0 mV a 700 mV con el sincronismo entre -300 mV y 0 V para una imagen de barras de color. También se ha representado la misma imagen para la señales en componentes. En este caso el margen dinámico de la señal de luminancia Y es desde 0 mV a 700 mV con el sincronismo igual al de RGB, y las señales diferencia de color Pb y Pr con el margen dinámico -350 mV a 350 mV (700mV en total).

⁴ NTSC: National Television System Committee, SMPTE: Society of Motion Picture and Television Engineer y EBU/PAL: European Broadcasting Union/Phase Alternating Line.

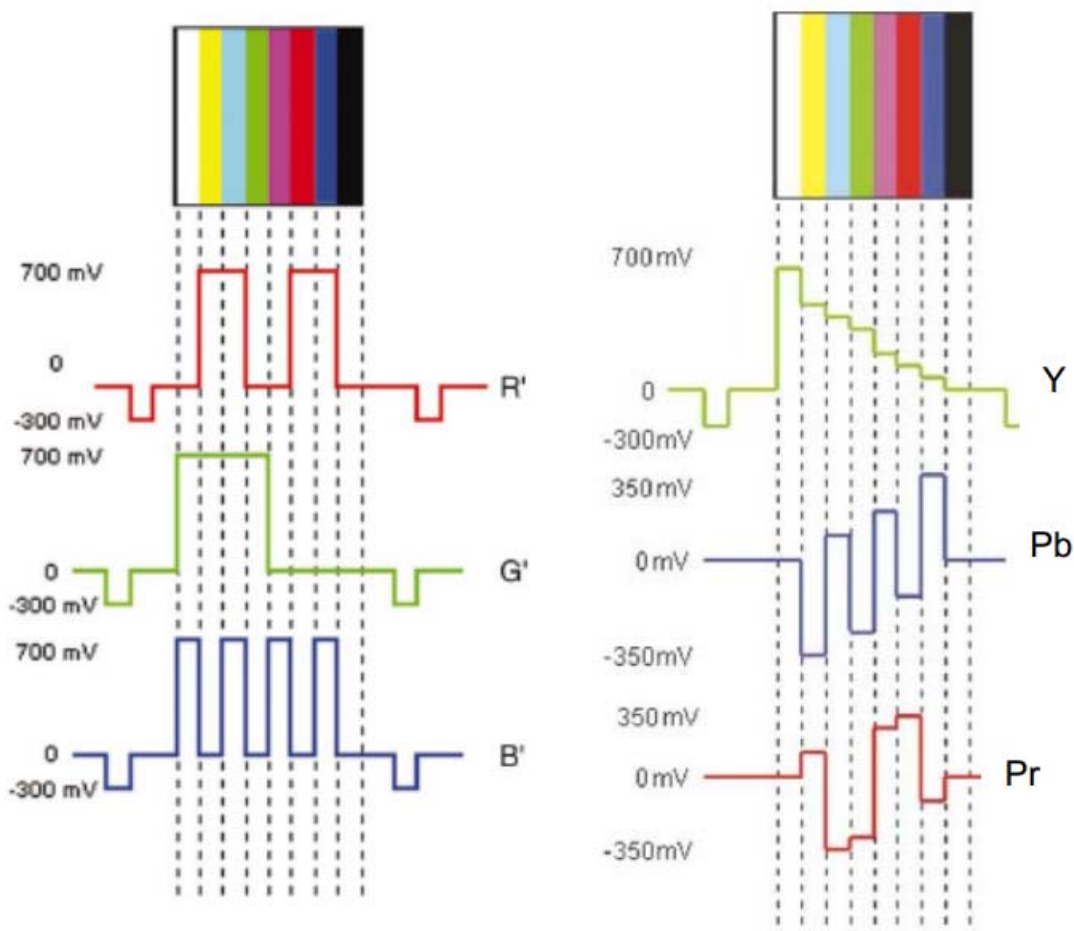


Figura 3-23. Señales RGB y Y Pb Pr de una imagen de barras de color.

3.3 Cámaras de Vídeo

Como ya se comentó al principio de este capítulo, las cámaras de vídeo captan una imagen a partir de la realidad circundante o escena y la convierten en señal eléctrica.

Así, la cámara se diseña para cumplir determinados aspectos:

- Frecuencia de cuadro, número de líneas, entrelazado o progresivo,...
- Conversión opto-electrónica: funciones colorimétricas en base a los colores primarios del sistema de vídeo, corrección gamma,...
- Amplitudes, anchos de banda, posible digitalización, información asociada de la imagen,...
- Correcciones y procesado de las señales para representar mejor la imagen captada.
- Interconexión con otros equipos.

Las cámaras se componen de varias partes que se montan entre sí:

1. Lente u objetivo: determina la imagen que se capta y la cantidad de luz que reciben los sensores.
2. Cuerpo de cámara: contiene el bloque óptico (filtros de temperatura de color, prisma dicróico y filtros ópticos asociados), los sensores y la electrónica para obtener y procesar la señal de vídeo.

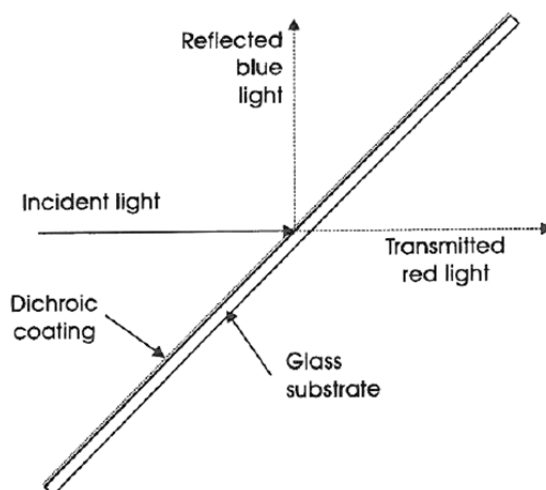


Figura 3-24. Filtro dicroico.

3. Adaptador de salida y posibles sistemas auxiliares.
4. Visor de imagen.
5. Soporte (trípode, pedestal, grúa, el operador de cámara,...).

3.3.1 Diagrama general de una cámara

El la Figura 3-25 se expone el esquema general de una cámara de vídeo.

-Preamplificadores de cámara: son los responsables de los ajustes finos de nivel y ganancia. No deben confundirse con los preamplificadores de los sensores que van integrados con ellos.

-Procesadores: se encargan de procesos generalmente no lineales de la señal de vídeo.

-Las señales RGB resultantes quedan disponibles a la salida de la cámara y también como entrada al codificador, encargado de obtener varios tipos de señal a partir de las RGB, como componentes $Y C_B C_R$, SDI, analógico, etc.

-El generador de sincros propio de la cámara genera las señales que controlan los diversos pasos. Este generador puede engancharse a un patrón exterior.

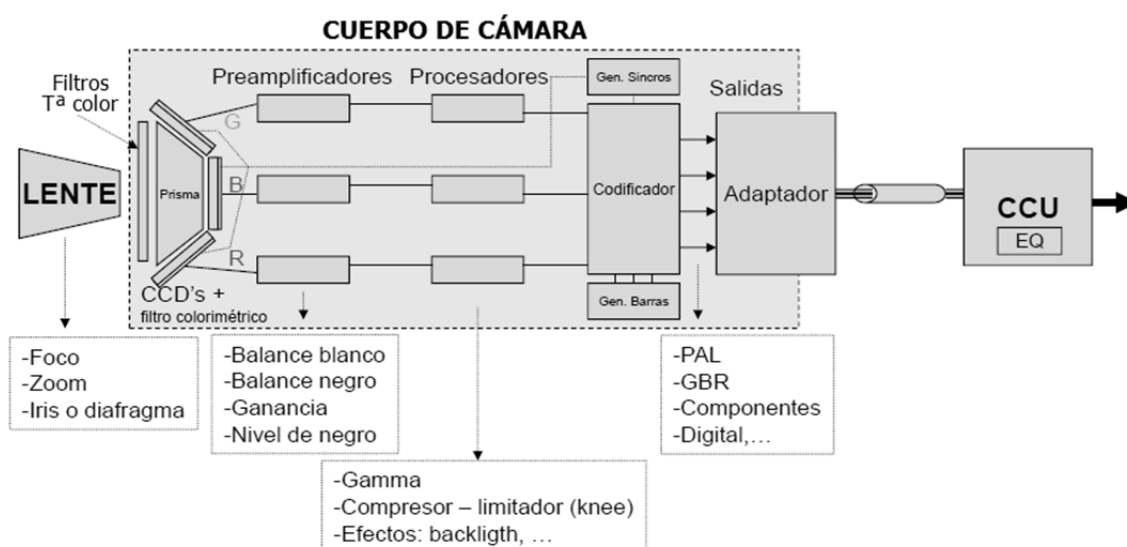


Figura 3-25. Esquema general de una cámara de vídeo.

Todas las señales quedan disponibles en el adaptador de salida, que en cámaras profesionales puede tener las siguientes opciones:

1. Salida a un equipo grabador portátil. Pueden ser grabadores incorporados en la cámara (cámaras CamCorder) o grabadores externos.
2. Salida a una Unidad de Control de Cámara (CCU: *Camera Control Unit*) con conexión de cable multipar o cable 'Triaxial'.

La salida proporciona, además de la señal de vídeo en sus distintos formatos, otras señales de posible uso simultáneo como: audio desde la cámara, señales de control remoto de la cámara, señales de control remoto del grabador, señal de intercomunicación, etc.

Todas las cámaras además incluyen un generador de señal de test (como mínimo unas barras de color) que se usan para algunos ajustes, principalmente relacionados con la compensación del efecto del cable de la cámara.

Los sensores son parte imprescindible de cualquier cámara. Prácticamente todas las cámaras de vídeo usan sensores CCD (*solid-state charge-coupled device*) o dispositivos en estado sólido de carga acoplada.

Más adelante se explicarán los diferentes tipos de sensores.

3.3.2 Bloque Óptico

La luz proveniente de una escena con la que somos capaces de ver, a través de una videocámara, lo que ocurre en dicha escena se denomina Iluminación. La fuente natural más común, el Sol, radia en un amplio espectro que no solo abarca la parte visible sino también rayos infrarrojo y ultravioleta. Sin embargo, las cámaras tienen que limitar su respuesta solo al espectro visible, que como se vio al principio del tema va de 400 nm a 700 nm de longitud de onda. Existen también muchas fuentes artificiales de iluminación, como lámparas incandescentes o luces fluorescentes. El ojo humano se intenta adaptar a las diferentes distribuciones espectrales hasta llegar a percibir a la fuente como un radiador de luz blanca.

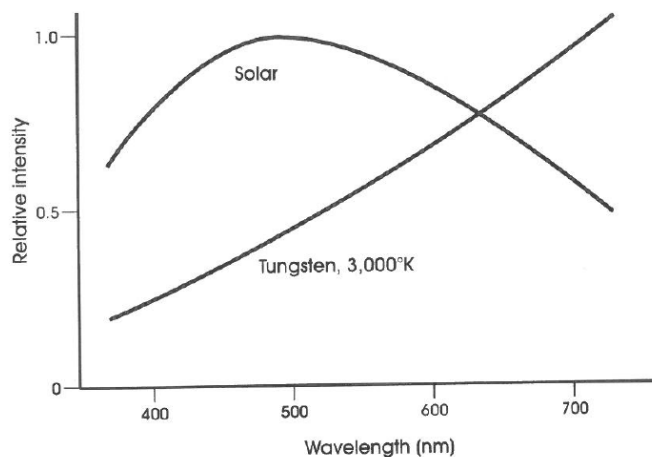


Figura 3-26. Distribuciones espectrales de la luz solar y el Tungsteno.

El sistema óptico completo consta de la lente y del bloque óptico. La lente, óptica u objetivo de las cámaras profesionales siempre es un elemento externo e intercambiable por otras lentes. El bloque óptico está constituido por una serie de elementos ópticos internos propios de cada cámara:

1. Filtros de temperatura de color. Ayudan a realizar el balance de blancos e interpretar adecuadamente los colores de la escena.
2. Filtros UV, de infrarrojos y paso bajo. Algunos de ellos son a veces implementados en la parte óptica de los sensores. El filtro paso bajo óptico reduce el efecto de *aliasing* que aparecerá en los sensores.
3. Prisma de separación de color o componentes RGB.

4. Filtros de ajuste colorimétrico.
5. Parte óptica de los sensores opto-electrónicos. En algunos dispositivos sensores, la superficie de los fotodiodos, se recubre con microlentes o con filtros de color para cada sensor.

3.3.3 Balance de blancos

Una cámara de vídeo no tiene una adaptación característica como la que tiene el ojo al percibir una escena lo cual debe ser proporcionado de manera electrónica. A esto se le denomina **balance de blancos**. En un sistema de RGB blanco corresponde con la igualdad de las tres componentes de la señal, en un sistema de Luma-Croma blanco corresponde al cero de la croma.

El balance de blancos en un sistema de visualización ha de ser ajustable y se debe establecer de forma que la señal de vídeo sea vista por el observador como blanco. Si la iluminación de la cámara cambia durante la grabación el balance de blancos tendrá que ser repetido o el observador notará la diferencia en los colores de la escena.

3.3.4 Cuantificación de la Iluminación

Las fuentes de luz pueden caracterizarse por su densidad de energía (vatios por centímetro cuadrado) recibidas en la escena, pero es más útil en término de iluminación definir una fuente por su intensidad luminosa, que representa la respuesta espectral del ojo. En la Figura 3-27 se muestra como se capta una imagen por la iluminación de un objeto o escena.

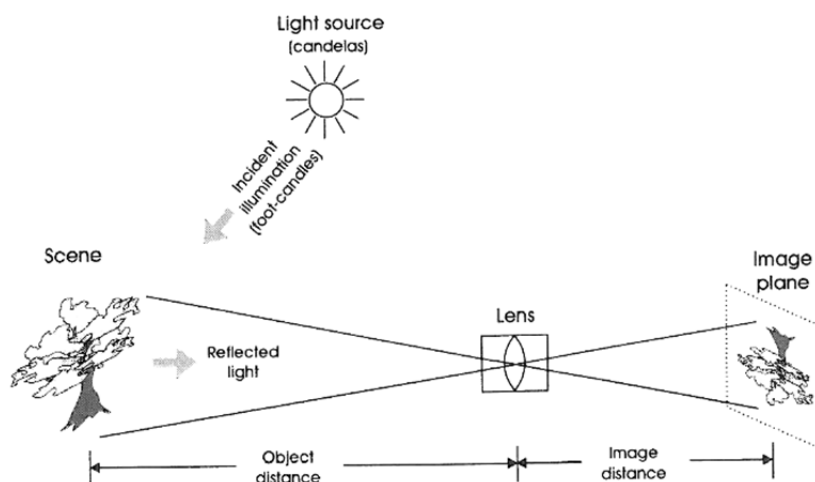


Figura 3-27. Iluminación

La luz es variación de energía a través del tiempo. La unidad de la variación de luminancia es el lumen y la intensidad de luminancia se expresa como *lumens* por unidad de ángulo sólido. Se usa el ángulo ya que la intensidad puede variar dependiendo del ángulo de incidencia de un rayo de la fuente respecto a la escena. La intensidad también varía con la distancia entre fuente y escena de manera inversamente proporcional a la inversa del cuadrado de la distancia entre ambas. Debido a ello es más descriptivo expresar la luz en una escena como *lumens* por metro cuadrado. Para ello se ha definido lux, que es lumen/m^2 . Una unidad equivalente es el *foot-candle*, que es lumen por pie cuadrado. La relación entre estas es $1\text{lux}=10,76\text{ foot-candles}$. La sensibilidad de las cámaras viene expresada generalmente en lux.

El brillo de una superficie es lo que ésta refleja de toda la iluminación que le llega. Dado que se trata de un fenómeno diferente al de la iluminación en sí tiene su propia unidad de medida: el nit que es una candela por metro cuadrado.

3.3.5 Sensibilidad de una Cámara

La sensibilidad de una cámara es la cantidad de luz que es necesaria para que los sensores den una señal de nivel adecuado (un valor concreto que dependerá de la normativa específica usada al hacer la medida). Se especifica con la siguiente definición: cantidad de luz, expresada en Lux, de la escena para que con un cierto número F, los sensores den un nivel de señal adecuado. El número F se calcula con la expresión (3-15):

$$F = \frac{f}{\Phi}. \quad (3-15)$$

Donde f es la distancia focal y Φ es el diámetro del diafragma. Así, por ejemplo, si una cámara tiene una sensibilidad de '2000 Lux a F4.0', dará una señal adecuada de una escena iluminada con 2000 Lux que sea captada por un sistema óptico con F4.0. Durante el manejo de las cámaras, el diafragma de la lente se ajusta a distintos números F según interese. En ese caso, la luminosidad del sistema óptico será distinta en cada momento. Si la lente se cierra a números F altos, la luminosidad del sistema óptico será muy cercana a dicho número F ajustado en la lente. Por eso es habitual que al hablar de sensibilidad de la cámara se indique la iluminación de escena necesaria para un cierto número F del diafragma de la lente.

3.3.6 Lentes

El sistema óptico completo consta de la lente y del bloque óptico. La lente, óptica u objetivo de las cámaras profesionales siempre es un elemento externo e intercambiable por otras lentes. El bloque óptico está constituido por una serie de elementos ópticos internos propios de cada cámara, representados de la Figura 3-28:

1. Filtros de temperatura de color. Ayudan a realizar el balance de blancos e interpretar adecuadamente los colores de la escena.
2. Filtros UV, de infrarrojos y paso bajo. Algunos de ellos son a veces implementados en la parte óptica de los sensores. El filtro paso bajo óptico reduce el efecto de *aliasing* que aparecerá en los sensores.
3. Prisma de separación de color o componentes RGB.
4. Filtros de ajuste colorimétrico.
5. Parte óptica de los sensores opto-electrónicos. En algunos dispositivos sensores, la superficie de los fotodiodos, se recubre con microlentes o con filtros de color para cada sensor.

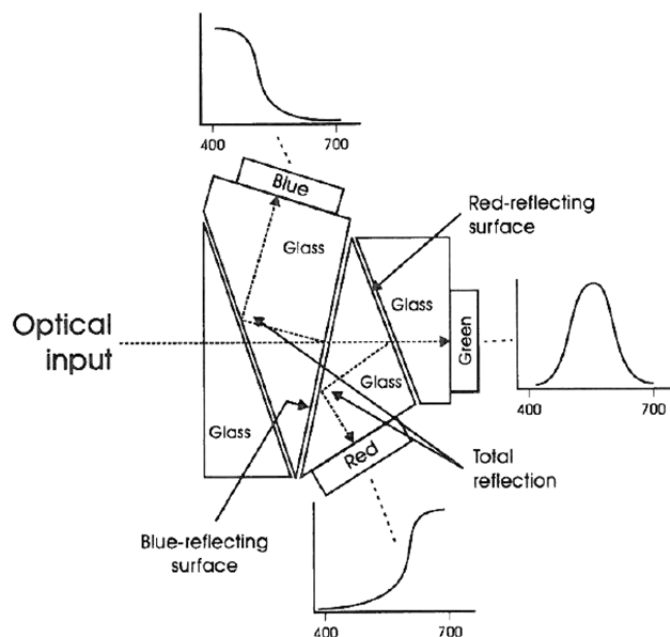


Figura 3-28. Detalle del bloque óptico de una cámara. El elemento más importante es el prisma separador, pero no es el único imprescindible para obtener la captación más correcta de la escena.

Una lente modela la visión de la cámara, determina el tamaño de las cosas, lo que cabe en el encuadre y qué objetos están en el primer plano y cuáles de fondo. Una lente para cámara de vídeo, al igual que para fotografía, se caracteriza por una serie de parámetros básicos que se definen de forma común, pero los valores dependen de cada aplicación. En muchas ocasiones, en vez de utilizar el término lente, se usa la palabra ‘objetivo’ u ‘óptica’. El objetivo es un conjunto de lentes centrado y convergente, agrupado en un sistema mecánico.

3.3.7 Tamaño del sensor y de la imagen

Aunque la lente sea circular, los sensores tienen una estructura rectangular que define la forma de las imágenes que vemos. El sensor rectangular queda inscrito en la circunferencia que forma la lente para aprovechar al máximo la luz que se capta, pero siempre se desperdiciará parte de lo que realmente se capta con el objetivo.

El diámetro del círculo de imagen se mide en pulgadas y sirve para especificar el tamaño de la imagen. Los tamaños más comunes son de 1/2", 2/3".

3.3.8 Distancia focal

La distancia focal es la distancia que existe entre el punto principal de la lente y el punto focal. Se mide en mm. La Figura 3-29 muestra que la distancia desde las lentes a la escena (distancia objeto) y la distancia desde las lentes hasta el plano de la imagen (imagen distancia) están geométricamente relacionadas. Sin embargo la imagen será enfocada en un único punto (o un pequeño rango de puntos) de acuerdo a la ecuación (3-16):

$$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f} \quad (3-16)$$

Donde s_o es la distancia objeto s_i es la distancia imagen y f es la longitud focal de la lente.

Si la distancia objeto es infinita la ecuación se reduce a $s_i=f$, es decir esta igualdad se cumple cuando la escena esta en el infinito. Para un determinado tamaño del plano imagen, como por ejemplo el tamaño de una superficie sensora, la distancia focal determinará el **ángulo de visión de la cámara**.

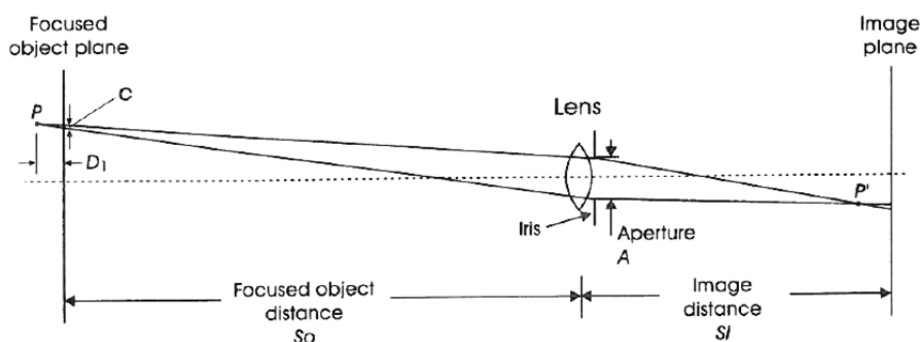


Figura 3-29. Distancia Focal.

El ángulo de visión se define como el margen angular que delimita los objetos que pueden ser enfocados en el sensor de la imagen. El ángulo de visión de una lente es el mismo en todas las direcciones al existir simetría de revolución respecto al eje. Sin embargo, dado que el sensor de la cámara no es redondo sino rectangular y con cierta relación de aspecto, la cantidad de imagen aprovechada en el sentido vertical es menor que en el horizontal y suele hablarse de ángulos efectivos de visión horizontal y vertical.

3.3.9 Apertura del diafragma

El brillo de la imagen viene dado por la cantidad de luz que pasa el objetivo e incide sobre el elemento sensor. Depende tanto de la luminosidad de la escena captada como de la lente. La lente dispone de un elemento que permite regular la luz que entra y por tanto, el brillo de la imagen. Este elemento es el **diafragma**, también

denominado iris, que puede bloquear el paso de la luz. La actuación del diafragma puede ser medida indicando el diámetro del orificio, llamado **diámetro del diafragma**. Este puede ser ajustado por el usuario entre varias posiciones como se muestra en la Figura 3-30.

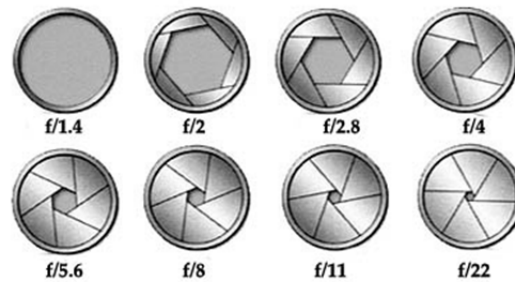


Figura 3-30. Obturador con diferentes aperturas (distintos números F).

3.3.10 Profundidad de campo

El rango de las distancias a las que se encuentran los objetos que quedan enfocados en una misma imagen se denomina profundidad de campo. Por ejemplo, una cierta lente enfocada a 5 metros puede dar imágenes enfocadas de objetos que se encuentren desde 4 metros hasta los 7 metros, luego la profundidad de campo es de 3 metros. Cuanto mayor sea la profundidad de campo más objetos se verán enfocados en la imagen. Si la profundidad de campo es muy pequeña, sólo se verá enfocado el objeto que está a distancia de enfoque y el resto de la imagen se verá desenfocada.

La profundidad de campo no es un valor fijo ni puede ajustarse directamente en la lente, sino que es el resultado de varios factores. En general se debe considerar que:

- Cuanto mayor es la distancia de enfoque, mayor es la profundidad de campo.
- Cuanto menor sea la apertura del diafragma mayor es la profundidad de campo.
- Cuanto menor sea la distancia focal de la lente, mayor es la profundidad de campo.

3.3.11 MTF

Cuando se quiere profundizar más sobre cuál es el detalle más pequeño de la imagen que puede captarse o sobre la apertura que es mejor usar en un momento dado, no basta con los conceptos manejados anteriormente, sino hay que recurrir a conceptos que evalúen la calidad de la transmisión de luz de la lente y no sólo su cantidad.

La MTF mide la capacidad de resolución de la lente de la siguiente manera:

Se capta con la lente una carta especial consistente en rayas blancas y negras verticales cuyo grosor y separación va disminuyendo desde el centro de la carta hacia los dos extremos.

Se analiza la señal eléctrica obtenida con sensores especiales, que consistirá en una sucesión de niveles altos (blancos) y bajos (negros). El contraste se calcula como la relación entre el nivel blanco y el negro.

La medida de MTF consiste en obtener una curva que representa el valor del contraste en función de la frecuencia espacial de las líneas de la carta. Dado que dicha frecuencia aumenta al separarnos del centro de la lente, se puede expresar la variación de MTF con dicha distancia en lugar de en función de valores concretos de la frecuencia espacial.

Llega un momento en que el contraste cae a 0 y es imposible distinguir una raya blanca de otra negra y lo que se ve es una zona gris: estamos por encima del denominado poder de resolución de la lente. La frecuencia espacial para la que ocurre esto nos indica cuán pequeño puede ser un detalle que queramos distinguir con dicho objetivo. En la Figura 3-31 se puede ver diferentes resultados de análisis de la MTF. La calidad del sensor de medida

usado debe ser superior a la de la lente para no afectar a la medida. No suelen usarse simples sensores de cámaras, cuyo MTF suele ser inferior al de la lente.

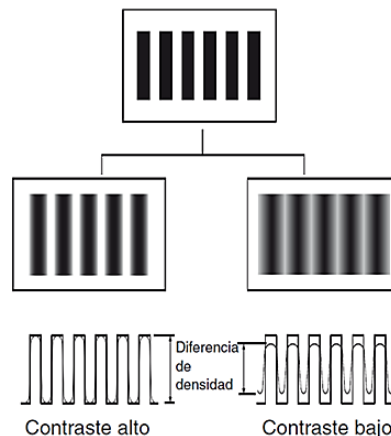


Figura 3-31. Diferentes resultados obtenidos del análisis de la MTF.

3.3.12 Sensores

Los sensores de la cámara se encargan de transformar la luz en señal eléctrica. Todos los sensores están basados en un array de elementos semiconductores fotosensibles (llamados foto-diodos) que generan una carga eléctrica proporcional a la luz que reciben. La diferencia entre ellos es cómo se extraen esas cargas hacia el conector de salida en el que aparece una señal de vídeo. El array que forman los sensores realiza un muestreo bidimensional de la imagen y cuando ésta tenga detalles de muy alta frecuencia espacial podrá aparecer *aliasing*, y no aparecerán representados.

Para eliminar el *aliasing* se toman dos medidas:

1. Usar un sensor con más píxeles de los necesarios para obtener una frecuencia de muestreo espacial mayor.
2. Usar un filtro paso bajo óptico para disminuir, en el MTF óptico, las componentes de la imagen que son de alta frecuencia espacial con respecto a la frecuencia de muestreo espacial del sensor. Al filtrar, la imagen perderá los detalles finos y resolución, pero se disminuye la posible aparición de *aliasing*.

Es habitual caracterizar el problema bidimensional de **muestreo por partes**, analizando dos muestreos, uno horizontal y otro vertical. Este análisis por partes no es completo pero es suficiente para definir bastantes aspectos.

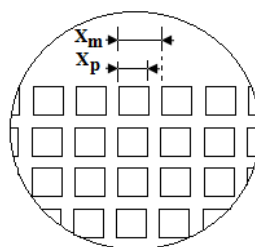


Figura 3-32. Esquema representativo de un array sensor con fotodiodos cuadrados. La separación entre ellos establece el muestreo espacial, X_m . El ancho del fotodiodo, X_p , establece la integración espacial, haciendo que el muestreo no sea puntual.

El análisis de la imagen, siguiendo líneas horizontales, permite definir las frecuencias horizontales y analizar un *aliasing* horizontal. Igual ocurre en vertical. La separación entre dos fotodiodos (píxeles del sensor) se mide en μm y es el espacio de muestreo. Su inverso es la frecuencia de muestreo espacial (horizontal o vertical) y se mide en μm^{-1} . La frecuencia de muestreo disminuye a medida que aumenta la distancia entre los píxeles del sensor.

La **resolución**, tanto en vertical como en horizontal, se puede definir en vídeo como el número de líneas blancas y negras que pueden distinguirse en el cuadrado central de la imagen.

Para aumentar la sensibilidad del sensor, cada fotodiodo tiende a ser lo mayor posible, por lo que su tamaño respecto a la separación entre ellos se acerca a 1.

Proceso de lectura de los sensores

Una vez los fotodiodos han adquirido una carga eléctrica proporcional al nivel de luz de la escena (o de la componente de color correspondiente), ésta debe extraerse o leerse para establecer una señal de vídeo a la salida del dispositivo sensor.

Para extraer las cargas es necesario:

1. Direccionar y acceder a la carga de un fotodiodo o fila de fotodiodos en cada momento. Son las funciones de escaneo del sensor, controladas por relojes específicos.
2. Extraer dicha carga llevándola desde cada fotodiodo al registro de salida.

Existen dos formas de extraer las cargas de los fotodiodos que dan lugar a dos familias de sensores: **CCD** y **CMOS** de la Figura 3-33.

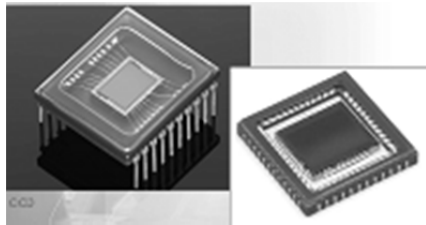


Figura 3-33. Sensor CCD (izqda.) y CMOS (dcha.).

CCD

Los **CCDs** (*Charge Coupled Device*) son básicamente registros de desplazamiento que mueven cargas y que incluyen los fotodiodos que las generan. Un sensor se configura agrupando CCDs verticalmente. Así, se pueden desplazar simultáneamente grupos horizontales de cargas correspondientes a líneas de imagen. Se usan técnicas de control digital, junto con métodos muy eficientes de electrónica de semiconductores, para mover la carga de un píxel a otro, para establecer y transferir un array bidimensional de cargas desde la parte sensora del CCD a la parte de registro de salida independiente del anterior. Al final, un registro de desplazamiento horizontal organiza el tren de cargas analógicas que van saliendo en serie del CCD convertidas en tensión. El proceso de extracción de cargas del sensor se conoce como 'mecanismo de transferencia' y existen tres mecanismos establecidos: *Frame transfer*, *Interline transfer* y *Frame-interline transfer*.

-Frame transfer (FT). En la Figura 3-34 se muestra un esquema simplificado de una estructura FT. Se usan dos arrays bidimensionales separados, uno como el captador de luz (el que crea las cargas) y el otro como almacén de las mismas (no expuesto a la luz).

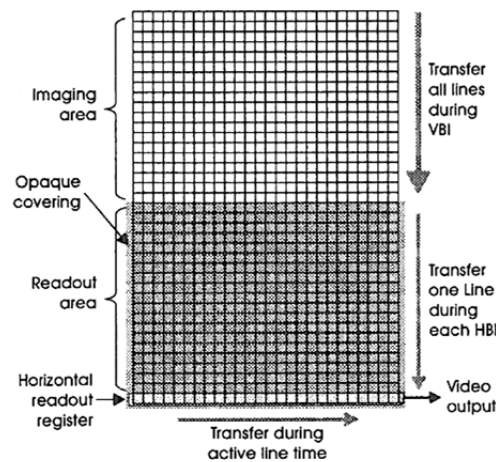


Figura 3-34. Arquitectura *Frame Transfer*.

Durante la parte activa de un campo la parte sensible se carga. Cuando empieza el borrado vertical, todo el array sensor se usa como un conjunto de registros de desplazamiento vertical para transferir el array de cargas al array de almacenamiento. Durante el siguiente periodo activo de campo, el array de almacenamiento se va leyendo a ritmo de vídeo, mientras que el array sensor se vuelve a cargar con la luz de la escena que se capta. El proceso descrito es para un sistema con entrelazado pero se puede hacer para cada cuadro de un sistema progresivo.

El CCD FT contiene sensores (fotodiodos) grandes y casi continuos. Tiene mucha sensibilidad y trasfiere fácilmente sus cargas. Respecto al muestreo, su apertura relativa puede considerarse igual a 1. Sin embargo, los sensores siguen siendo estimulados por la luz durante el proceso de transferencia (salvo que se usara un obturador mecánico, lo que no es muy habitual), haciendo que las cargas sean modificadas según van pasando de una celda sensible de CCD a otra. En esta estructura, cada píxel es tanto un sensor como una celda de desplazamiento, provocando un *smear* vertical.

El efecto *smear* se produce cuando algún elemento de la imagen tiene luminosidad muy intensa (un foco de un coche, el sol, etc.) y aparece una línea vertical no existente en la escena y centrada en el elemento son sobreluz. Este problema se puede solucionar únicamente con un obturador mecánico no muy usado en cámaras de vídeo, lo que hace del CCD FT un dispositivo de poco uso y justifica el diseño de otras estructuras CCD.

-Interline transfer (IT). En la Figura 3-35 se puede ver su esquema simplificado. Difiere del CCD FT en que sólo tiene un array bidimensional, que es el sensible, y que el registro de desplazamiento CCD está separado de las celdas sensibles (fotodiodos), dispuesto como un canal siguiendo la vertical del sensor, entre los fotodiodos contiguos en horizontal. Esto reduce el tamaño de la parte sensible del píxel resultando una menor sensibilidad. También supone una menor apertura relativa en cuanto al efecto en el muestreo de la imagen.

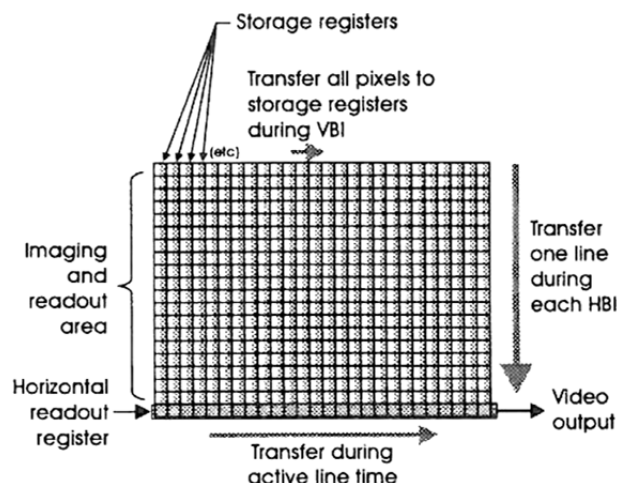


Figura 3-35. Arquitectura *Interline Transfer*.

Durante el periodo de campo activo, los fotodiodos van acumulando carga. En el periodo de borrado vertical que sigue, las cargas se transfieren casi instantáneamente al registro de desplazamiento CCD vertical adyacente, que no son sensibles a la luz y no alteran las cargas que pasan por ellos. Durante el siguiente periodo activo de campo, mientras los fotodiodos vuelven a generar carga en cada celda sensible, los registros verticales van descargando línea a línea las cargas del campo anterior en un registro de salida horizontal, del que va saliendo la señal de vídeo.

Los relojes de desplazamiento vertical son ahora más lentos que en el FT. Los canales CCD verticales dan lugar al nombre de *interline*, aunque parecería más lógico el nombre de *intercolumnn*.

El *smear* no debería existir y no es necesario proponer un obturador mecánico. Sin embargo, sobreluces mucho más altas pueden provocar un desbordamiento de carga desde el fotodiodo al canal de transferencia provocando nuevamente la aparición de dicho efecto.

-Frame-Interline transfer (FIT). Es un sistema mezcla de los dos anteriores y su principal aportación es la posibilidad de implementar la obturación electrónica de velocidad variable. Esto permite obtener señales de vídeo cuyos cuadros contengan imágenes claras, no borrosas y casi congeladas de objetos en movimiento rápido. Como se aprecia en la Figura 3-36, el CCD FIT tiene dos arrays bidimensionales, uno contiene los fotodiodos y los canales de desplazamiento verticales inter-líneas y el otro es el de almacenamiento, desde el que se hace la lectura a ritmo de vídeo a través del registro horizontal de salida. Durante el periodo activo, los fotodiodos cargan las celdas sensibles a la luz y el array bidimensional de almacenamiento se lee. Durante el periodo de borrado siguiente, se realizan dos procesos de transferencia consecutivos a alta velocidad:

1. Transferencia casi instantánea de las cargas al registro vertical interlínea.
2. Transferencia a alta velocidad al array de almacenamiento.

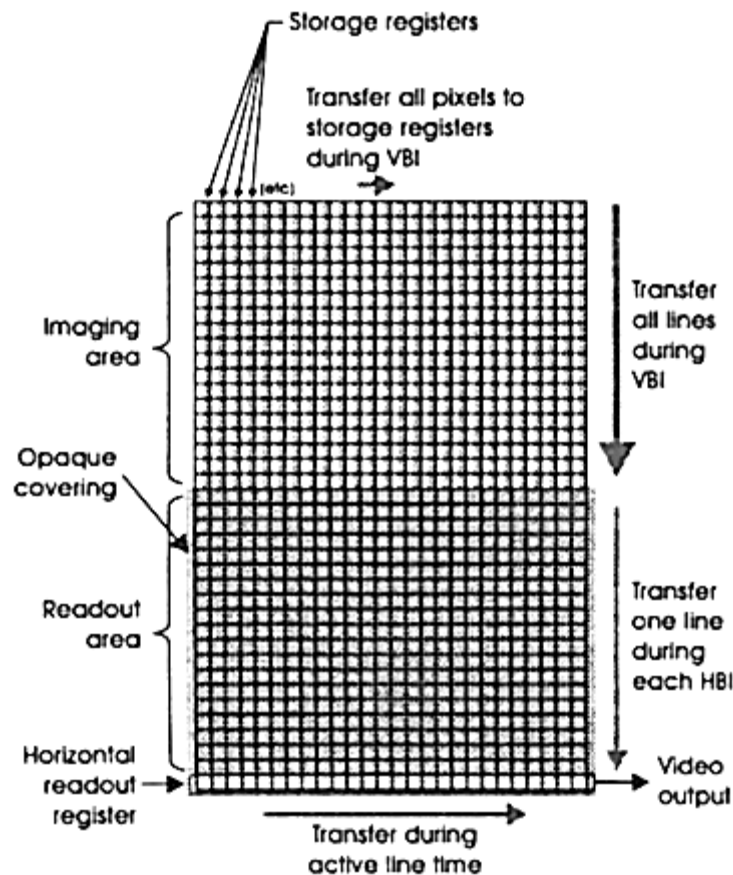


Figura 3-36. Arquitectura *Frame-Interline Transfer*.

Al hacerse la transferencia vertical en poco tiempo, y no durante todo un periodo activo como en CCD-IT, hay menos posibilidad de que una sobreluz muy alta desborde un fotodiodo afectando al canal y provocando *smear*. Por tanto, el *smear* que puede llegar a aparecer en un CCD-FIT es extremadamente bajo.

CMOS

Los **CMOS** (*Complementary Metal Oxide Semiconductor*) no se basan en registros de desplazamiento para sacar las cargas, sino en la misma técnica usada en chips de memoria CMOS. Cada celda de memoria consiste en un fotodiodo que genera la carga y en lugar de contener un bit (nivel alto o nivel bajo) contiene la carga concreta. Cuando es el momento de extraer o leer esa celda o píxel, un transistor CMOS asociado a la celda la conecta con un canal conductor (línea de bus) que la lleva directamente a un registro de salida. Normalmente los buses se integran en el chip estableciendo una línea para cada vertical de imagen y se puede transferir de golpe toda una línea de imagen a un registro horizontal de salida, como se hace en los CCDs. En la Figura 3-37 se muestra la estructura de estos sensores.

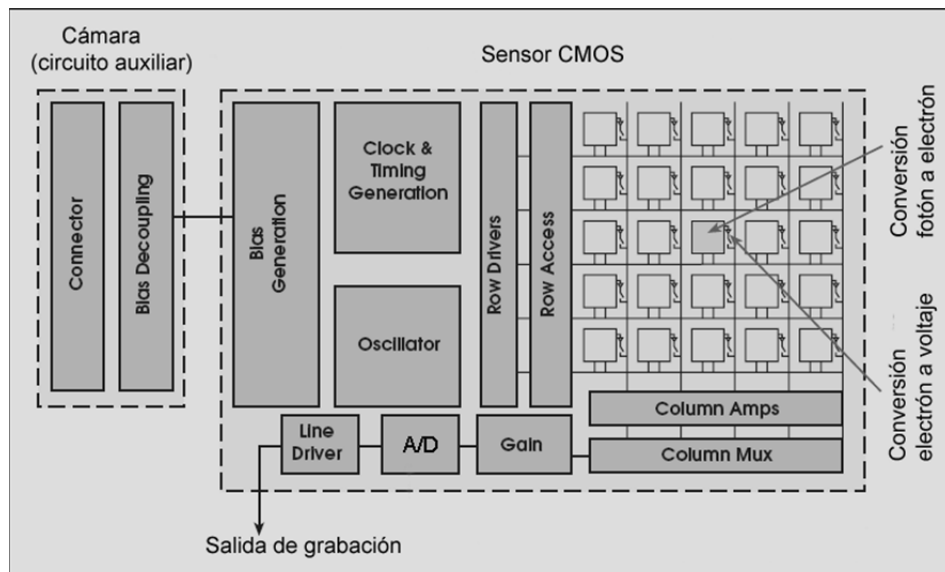


Figura 3-37. Un sensor CMOS es un array de fotodiodos que llevan asociados una celda CMOS de almacenamiento. En cada momento, la circuitería de direccionamiento o selección de filas-columnas, activa el transistor CMOS de la celda correspondiente para que la carga salga por los hilos conductores hacia el amplificador de salida.

El sensor CMOS y la matriz de Bayer

Las principales diferencias entre los CCD y los CMOS se podrían resumir en los siguientes puntos:

- Los CCD son menos ruidosos que los CMOS porque son más sensibles a la luz.
- Los CCD son menos sensibles a los colores que los CMOS.
- Los CCD consumen más energía que los CMOS.
- Los CCD son más caros de fabricar que los CMOS.

Actualmente se están desarrollando sensores CMOS cada vez mejores por lo que, si bien antes se solían encontrar sensores CCD en cámaras de alta gama y CMOS en gamas inferiores, ahora casi todas las cámaras que salen al mercado incorporan sensores tipo CMOS.

A pesar de la evolución y abaratamiento de los sensores CMOS, se han diseñado estrategias para reducir aún más los costes de las cámaras. La principal estrategia consiste en anteponer al sensor un filtro de **Bayer**.

El filtro Bayer consiste en una matriz de colores rojo, verde y azul ordenada en el espacio de forma regular. Se utilizan el doble de captadores verdes porque el ojo humano es más sensible al verde que a otros. De ese modo no se captan los tres colores en cada píxel de la imagen, sino uno sólo en cada uno de ellos y se aumenta la resolución del sensor sin aumentar su precio. El inconveniente es que es necesario interpolar la información resultante para rellenar la información de color que falta.

Además, delante del filtro de Bayer se antepone otro filtro denominado filtro de antialiasing (*AA filter*) que básicamente elimina las frecuencias más altas de la imagen, es decir, que emborrona los detalles que son más finos que el propio filtro de Bayer, con el fin de reducir los artefactos de aliasing en la imagen captada. Generalmente también se incluye un filtro infrarrojo, *IR Cut Filter* (ICF), para reducir la sensibilidad del conjunto sensor al infrarrojo (los CCD son fruto de la investigación militar y se diseñaron originalmente para ser muy sensibles al infrarrojo y poder usarse en visión nocturna).

Existe cierta controversia sobre la conveniencia de tener un filtro AA antepuesto al sensor. Hasta el punto de que hay usuarios que eliminan el filtro de sus cámaras (operación nada recomendable para novatos) para ganar algo de resolución o aumentar la sensibilidad de la cámara en fotografía infrarroja. Aquí podéis ver una interesante comparativa entre tener y no tener el filtro en una Nikon D200.

3.4 Tratamiento de las señales RGB

Las señales obtenidas directamente de los sensores no cumplen algunos de los requisitos para ser la mejor representación de la escena captada ni para ser correctamente interpretadas por otros equipos a los que se conecten. Por eso es necesario el tratamiento de dichas señales antes de ser usadas a la salida de la cámara.

Preamplificadores

Los preamplificadores de RGB se encargan de ajustar el nivel de componente continua y la ganancia para cada una de las señales.

Procesadores

Los preamplificadores sólo realizan funciones que supongan cambios de ganancias y niveles de continua. El resto de las funciones deben hacerse con circuitería o algoritmos especiales dando lugar a los denominados **procesadores**. Algunas de las funciones más habituales de los procesadores son: Corrección de Gamma, compresor-limitador, corrector de no uniformidad de luz, corrector colorimétrico o realzador de bordes.

Codificadores y adaptador de salida

Las señales RGB sólo suelen ser usadas en aplicaciones específicas, por ejemplo para hacer un *chroma-key* de alta calidad. Cuando salen de la cámara, están al principio de la cadena y estas señales RGB tienen más calidad que las señales codificadas a partir de ellas.

Lo más habitual es que se quiera usar la señal de cámara en formato componentes tanto analógicos como digitales. El codificador realiza la transformación RGB a $YC_B C_R$, ya sea de forma analógica o digital. En el caso de la transformación digital hay un corte en el ancho de banda y por tanto en la resolución de la imagen de la cámara.

Para usos sencillos sigue siendo normal que la cámara incluya codificadores a PAL, los cuales pueden realizar una transformación de forma analógica o digital, para al final dar una señal analógica.

Las señales obtenidas por el codificador quedan disponibles para su conexión al exterior de la cámara a través del adaptador de salida. En el caso de adaptador de salida para cable o manguera multipar, cada señal se asocia a unos contactos y a unos cables diferentes de dicha manguera multipar de conexión. Si el adaptador es para cable triaxial, el adaptador incorpora un dispositivo con circuitería de modulación en frecuencia. Además de las conexiones de salida indicadas antes, usadas para conectar la cámara a una CCU o a un equipo grabador, el adaptador de salida dispone de otros conectores para usar las señales en otras funciones.

El cable triaxial equivale a dos cables coaxiales, para transmitir la señal en cada sentido. Para cada sentido debe realizarse un multiplex por división de frecuencia de las señales a enviar, usando moduladores.

Unidades de control de cámara: CCU (*Camera Control Unit*)

Cuando la cámara se usa en estudio o conectada a una unidad móvil, siempre se conecta al resto de la instalación a través de una unidad de control de cámara o CCU cuyos objetivos son:

1. Permitir a un operador técnico modificar la mayoría de los parámetros ajustables de la cámara ya sea en la fase de ajustes previos a la toma o durante la toma de la escena.
2. Servir de interface entre la cámara y el resto de la instalación.

La CCU envía a la cámara las señales de control remoto para que la cámara efectúe los ajustes que se piden desde el panel de control de la CCU: ajuste de ganancias de preamplificadores, ajuste de detalle, ajuste de compresor, activación del generador de barras de color, etc.

En estudios con varias cámaras es habitual que todas sean iguales y que se quieran ajustar de igual manera. Para ello existe la opción de copiar los ajustes hechos en una CCU al resto de las CCUs. Suele existir un equipo de centralización de CCUs que recibe generalmente el nombre de MCU (*Master Control Unit*).

3.5 Cámaras analógicas y digitales

Aunque el sensor tipo CCD o el CMOS son dispositivos que discretizan la imagen, la señal que dan a su salida es continua, es decir, analógica. Si la cámara es totalmente digital, la señal de salida de los sensores es digitalizada desde el principio y todos los procesos descritos anteriormente (balances, ganancia, procesado no lineal, etc.) se realizan digitalmente, y solo al final las salidas analógicas son obtenidas por conversores D/A.

Algunas cámaras son sólo parcialmente digitales, pues sólo algunos de sus procesos se realizan tras digitalizar la señal. Inicialmente las cámaras hacían todos los procesos con circuitería analógica, incluso los no-lineales, pero la tecnología actual permite la total digitalización de las cámaras.

La digitalización para el procesado de las señales RGB de las cámaras, tanto procesado lineal como no-lineal, requiere el uso de 14 bits por muestra, con cuantificación lineal, en cada uno de los tres canales RGB. La frecuencia de muestreo es menos repetida entre fabricantes, eligiendo cada uno la que mejor se adapta al MTF y la resolución que quiere obtener de la cámara. Al final, para la señal SDI de salida, se re-muestreará la señal digital para tener los valores establecidos para interconexión.

La influencia de los mercados semi-profesionales y de consumo ha hecho que se extienda la idea no profesional de que una cámara digital es aquella que da señal digital (SDI o comprimida) o que graba sobre un formato digital. El mercado de la fotografía ha ido más allá al hacer entender que una cámara con sensor discreto es una cámara digital. Debe tenerse cuidado con este tipo de confusiones.

3.6 Cámaras con soporte de grabación: Camcorders

Las cámaras que integran un soporte de grabación suelen tener todos los elementos de cualquier cámara, incluidos el codificador y el adaptador de salida, aunque tengan menos opciones disponibles al usuario. Además, el formato de señal más útil para el formato del soporte de grabación se hace llegar internamente a este para que realice dicha grabación. Pueden considerarse como dos equipos independientes que se han montado en una carcasa común y se encuentran conectados internamente. En la Figura 3-38 aparece una *Camcorder JVC*.



Figura 3-38. Cámara de vídeo tipo *Camcorder*.

3.7 Funcionamiento de una cámara de vídeo

Podemos explicar su funcionamiento por pasos. Primero, la luz pasa por la óptica de la cámara que consiste en un **sistema de lentes** que enfocan la escena adecuadamente y pueden aplicar un zoom para acercar o alejar la escena, y de un **diafragma o iris** que ajusta la cantidad de luz que pasa. La luz que proviene de la óptica pasa por el **sistema sensor**, compuesto por un filtro de temperatura de color para compensar el espectro de la luz de iluminación y un prisma dicróico que descompone la luz en las tres componentes básicas que se utilizan en vídeo: el rojo (R o *red*), el verde (G o *green*) y el azul (B o *blue*). Justo en la otra cara de cada lado del prisma están los captadores o sensores, actualmente dispositivos CCD. (A diferencia de una cámara fotográfica, aquí los CCD son 3, uno para cada color). Entre el prisma y los sensores hay unos filtros de ajuste colorimétrico según los colores primarios del sistema. Esta imagen leída por los CCD y su sistema de muestreo es conducida luego a los **circuitos preamplificadores** en los que se realiza el ajuste de niveles (blanco, negro), de balances (blanco y negro) y de ganancias. De los preamplificadores las señales se encaminan a los **procesadores**, donde se realizarán las correcciones de gamma, limitadores (*knee*) y efectos tipo contraluz (*backlight*), máscara (*masking*), etc. Casi todos ellos son procesos no-lineales y sofisticados. Después, las señales ya están listas para salir al sistema de producción o para ser grabadas. Se envían también a los circuitos de visionado, los cuales muestran la

imagen en el visor de la cámara y se transmiten mediante los correspondientes conectores de salida. La **salida** básica es RGB y/o Componentes ($Y C_B C_R$), aunque no siempre están disponibles. También se tiene, aun hoy en día, la salida del sistema analógico de TV: PAL, por lo que el **codificador** está presente en todas las cámaras. Añadido al mismo estará el codificador de señal digital SDI (ITU.R- 601/656). Estas señales son mandadas mediante el **adaptador triaxial** o el correspondiente cable a la CCU, que se encargará de encaminarlas al sistema de producción al que pertenece la cámara. Los sistemas auxiliares de comunicación **intercom** y los sistemas de control de la óptica y del **tally** residen en circuitos electrónicos en una placa auxiliar. El Tally es una luz roja que se activa en la cámara cuando la cámara está siendo utilizada en el control para ser grabada o emitida.

Todo ello es alimentado por la fuente de alimentación que se encarga de generar las diferentes tensiones de alimentación necesarias para los circuitos electrónicos y ópticos. Estas tensiones suelen partir de una única tensión de alimentación. No obstante, las cámaras de estudio funcionan con tensión de red, mientras que las ENG (cámaras de reporteros) pueden ser alimentadas por baterías de 12 V. Además de la circuitería indicada, la cámara incluye un **generador de señal de test** (típicamente barras de color). Muy importante es el **circuito generador de sincros** y relojes, cuyas señales controlan los procesos de la cámara e insertan las señales de sincros en la señal de vídeo de salida. Este generador puede funcionar por libre o engancharse a la señal del generador de sincronización patrón del estudio.

3.8 Cuestionario

1-El número F (*F-stop*) permite evaluar la cantidad de luz que pasa por la lente teniendo en cuenta el diámetro del diafragma y la distancia focal.

- a) Verdadero.
- b) Falso.

2-La distancia focal:

- a) Depende tanto de la luminosidad de la escena captada como de la lente.
- b) Es la distancia a la escena que se quiere captar
- c) Se define como el margen angular que delimita los objetos que pueden ser enfocados en el sensor de la imagen.
- d) Determina el ángulo de visión de la lente.

3-El espectro visible abarca:

- a) Todo el espectro electromagnético.
- b) De 380 Hz hasta 720 Hz.
- c) De 780 nm hasta 380 nm.
- d) Ninguna de las anteriores es cierta.

4-Asociar las siglas a su definición.

- | | |
|---------|---|
| a) ISV. | 1-Mínima distancia hasta un objeto para ser enfocado. |
| b) ISH. | 2-Línea activa de vídeo digital. |
| c) LAD. | 3-Impulso de sincronismo horizontal. |
| d) MOD. | 4-Interfaz digital serie. |
| e) MTF. | 5-Dispositivos de carga acoplada. |
| f) EAV. | 6-Início de video activo. |
| g) SDI. | 7-Fin de vídeo activo. |
| h) SAV. | 8-Impulso de sincronismo vertical. |
| i) CCD. | 9-Función de transferencia de modulación. |

5-Asociar cada dispositivo con sus funciones principales.

- | | |
|-----------------------|------------------------------|
| a) Preamplificadores. | 1-Efectos. |
| b) Procesadores. | 2-Control de ganancia. |
| | 3-Compresión y limitación. |
| | 4-Balance de blanco y negro. |
| | 5-Corrección Gamma. |

6- La mezcla aditiva espacial es una característica del sistema visual humano por la que dos puntos diferenciados son percibidos como uno solo si el ángulo que estos forman con respecto al ojo es menor a un minuto de grado.

- a) Verdadero.
- b) Falso.

7- Para que exista una coordinación entre lo captado y su reproducción:

- a) Los sistemas de vídeo incorporan señales de sincronismo: ISH y ISV.
- b) Se diseñan los sistemas de vídeo para que tengan auto-sincronización.
- c) Es necesario sistemas auxiliares basados en memorias que provocan un *frame* de retardo.

d) Ninguna de las anteriores es correcta.

8- Calcular en el sistema de televisión europeo (25 imágenes por segundo) el periodo de una línea de vídeo.

9- Los preamplificadores de RGB de las cámaras:

- a) Sirven de interface entre la cámara y el resto de la instalación.
- b) Suelen ser usadas en aplicaciones específicas, por ejemplo para hacer un *chroma-key* de alta calidad.
- c) Sólo realizan funciones que supongan cambios de ganancias y niveles de continua.
- d) Se encargan de ajustar el nivel de componente continua y la ganancia para cada una de las señales.

10- La MTF evalúen la cantidad de la transmisión de luz de la lente y no sólo su calidad.

- a) Verdadero.
- b) Falso.

11-La compatibilidad directa significa:

- a) Que los sistemas monocromo puedan reproducir señales en color.
- b) Que los sistemas de color puedan reproducir la señal monocroma.
- c) Que los sistemas analógicos reproduzcan señales digitales.
- d) Ninguna de las anteriores es correcta.

12-La compatibilidad inversa significa:

- a) Que los sistemas monocromo puedan reproducir señales en color.
- b) Que los sistemas de color puedan reproducir la señal monocroma.
- c) Que los sistemas analógicos reproduzcan señales digitales.
- d) Ninguna de las anteriores es correcta.

13-El margen dinámico de la señal de luminancia Y es:

- a) De 0V a 350 V.
- b) De -350 mV a 350 mV.
- c) De 0 V a 700 V.
- d) Ninguna de las anteriores es cierta.

14-Las señales diferencia de color tiene un margen dinámico de:

- a) De 0V a 350 V.
- b) De -350 mV a 350 mV.
- c) De 0 V a 700 V.
- d) Ninguna de las anteriores es cierta.

15-La corrección gamma aplicada sobre los sistemas de vídeo:

- a) Contrarresta el efecto de la MTF sobre las frecuencias altas, de modo que la distancia focal toma el valor inicial indicado por el fabricante.
- b) Es usada por dispositivos de reproducción de vídeo profesional debido a la necesidad de altísimas resoluciones y bajas tasas de ruido.
- c) Corrige el comportamiento no lineal de algunos elementos que componen la cadena de reproducción.
- d) Ninguna de las anteriores es cierta.

16-La ley de Ferry-Porter establece una relación el valor de brillo (B) se una imagen en pantalla y la frecuencia de cuadro crítica por debajo de la cual se aprecia parpadeo es f_c .

- a) Verdadero.
- b) Falso.

17-Las unidades de control de cámara o CCU:

- a) Permite a un operador técnico modificar la mayoría de los parámetros ajustables de la cámara ya sea en la fase de ajustes previos a la toma o durante la toma de la escena.
- b) Equivale a dos cables coaxiales, y sirve para transmitir la señal en cada sentido.
- c) Sirve de interface entre la cámara y el resto de la instalación.
- d) Las respuestas a) y c) son correctas.

18-La arquitectura *Frame Transfer* de los sensores CCD:

- a) Es un sistema mezcla entre las arquitecturas FIT e IT.
- b) Tiene un array bidimensional, que es el sensible, y que el registro de desplazamiento CCD está separado de las celdas sensibles (fotodiodos), dispuesto como un canal siguiendo la vertical del sensor, entre los fotodiodos contiguos en horizontal.
- c) Usa arrays bidimensionales separados, uno como el captador de luz (el que crea las cargas) y el otro como almacén de las mismas (no expuesto a la luz).
- d) Ninguna de las anteriores es cierta.

19- El prisma dicroico separa los rayos luminosos en colores específicos: generalmente los componentes rojo, azul y verde para poder así almacenar cada uno en su respectivo CCD (dispositivo de carga acoplada).

- a) Verdadero.
- b) Falso.

20- Los *Camcorders* son cámaras que integran un soporte de grabación suelen tener todos los elementos de cualquier cámara, incluidos el codificador y el adaptador de salida.

- a) Verdadero.
- b) Falso.

4 MONITORES Y PROYECTORES

Un **monitor de vídeo** es un dispositivo electrónico que permite visionar señales de vídeo. Han de ser fiables para estar mucho tiempo funcionando (a veces siempre encendidos) y robustos para soportar el trasiego de una producción o estar instalados en una unidad móvil.



Figura 4-1. Monitor profesional Panasonic.

Un **proyector de vídeo** o **vídeo proyector** es un aparato que recibe una señal de vídeo y proyecta la imagen correspondiente en una pantalla de proyección usando un sistema de lentes, permitiendo así mostrar imágenes fijas o en movimiento.

Todos los proyectores de vídeo utilizan una luz muy brillante para proyectar la imagen, y los más modernos pueden corregir curvas, borrones y otras inconsistencias a través de los ajustes manuales. Los proyectores de vídeo son mayoritariamente usados en salas de presentaciones o conferencias, en aulas docentes, aunque también se pueden encontrar aplicaciones para cine en casa. La señal de vídeo de entrada puede provenir de diferentes fuentes, como un sintonizador de televisión (terrestre o vía satélite), un ordenador personal,...

Otro término parecido a proyector de vídeo es retroproyector el cual, a diferencia del primero, se encuentra implantado internamente en el aparato de presentación y proyecta la imagen desde la parte de atrás.



Figura 4-2. Proyector de vídeo.

4.1 Conceptos básicos

Brillo

El brillo determina el nivel de intensidad de iluminación de fondo en la imagen presentada en pantalla. A mayor tamaño de la pantalla mayor iluminación será necesaria para producir un determinado nivel de brillo. La unidad de medida del brillo es el **nit** o candela por metro cuadrado (cd/m^2). Se recomienda 300 cd/m^2 para monitores que se usen en interiores y de 1000 cd/m^2 para el uso de estos en exteriores. Una denominación más específica del brillo es la esterancia luminosa.

Contraste

Se define el contraste como la diferencia relativa de intensidad luminosa entre los puntos de una imagen. Marca la diferencia de intensidad entre los elementos blancos y negros representados en la pantalla, es decir, es la proporción de brillo entre un píxel negro y uno blanco que el monitor es capaz de reproducir. Si el ambiente está

muy iluminado o hay reflejos en la pantalla, el contraste baja. Se usan valores de contraste con una relación de 500:1 o 5000:1.

-Contraste dinámico

El contraste dinámico es una nueva tecnología que se ha incorporado en el último tiempo en las pantallas LCD, no así en las pantallas basadas en LEDs o Plasma (ya que estas últimas tienen un contraste estático excelente). Es una técnica que modifica intencionalmente el brillo en determinadas imágenes para tratar de optimizar artificialmente el contraste. La técnica funciona analizando cada píxel de la imagen y bajando la iluminación para las imágenes más oscuras, ajustando así automáticamente el contraste en cada una. Por ejemplo, en una imagen donde predominan las escenas oscuras bajará la iluminación para que el negro sea más negro. El contraste dinámico no siempre funciona de manera adecuada y es un dato más fiable el proporcionado para el contraste estático. Si los valores del contraste estático superan los 1000:1, el monitor es de buena calidad.

Resolución

Hay que diferenciar dos conceptos: resolución de pantalla y resolución de imagen.

-Resolución en pantalla. Cantidad de puntos en sentido horizontal y vertical que se pueden presentar en pantalla (número de puntos por unidad de longitud). Como ya se comentó en el tema anterior, la resolución depende del ancho de banda de vídeo, de las frecuencias de barrido horizontal y vertical y del tamaño del punto luminoso en pantalla. Los monitores planos tienen varias opciones de resolución 800x600 píxeles, 1280x720 píxeles, 1920X1080 píxeles, etc. Pero solo es en una de estas opciones en las que pueden reproducir imágenes fielmente, que suele ser generalmente la más alta.

-Resolución de imagen. Indica cuanto detalle puede observarse en ésta. El término es comúnmente utilizado en relación a imágenes de fotografía digital, pero también se utiliza para describir cuan nítida (como antónimo de granular) es una imagen de fotografía convencional (o fotografía química). Tener mayor resolución se traduce en obtener una imagen con más detalle o calidad visual. Una imagen con poca resolución aparecerá con esta carencia por muy buena resolución que tenga el monitor donde se representa.

En la Figura 4-3 se puede ver una comparación de las resoluciones entre diferentes formatos de vídeo. A efectos de la resolución, también hay que destacar el **tamaño del punto**, representado en la Figura 4-4, que corresponde a la distancia que hay entre dos luminóforos (subpíxeles) del mismo color. Es un parámetro del que dependerá la nitidez de la imagen: si el tamaño del punto es muy grande, la imagen se verá granulada. Generalmente los monitores tienen tamaños de punto menores a 0,28 mm.

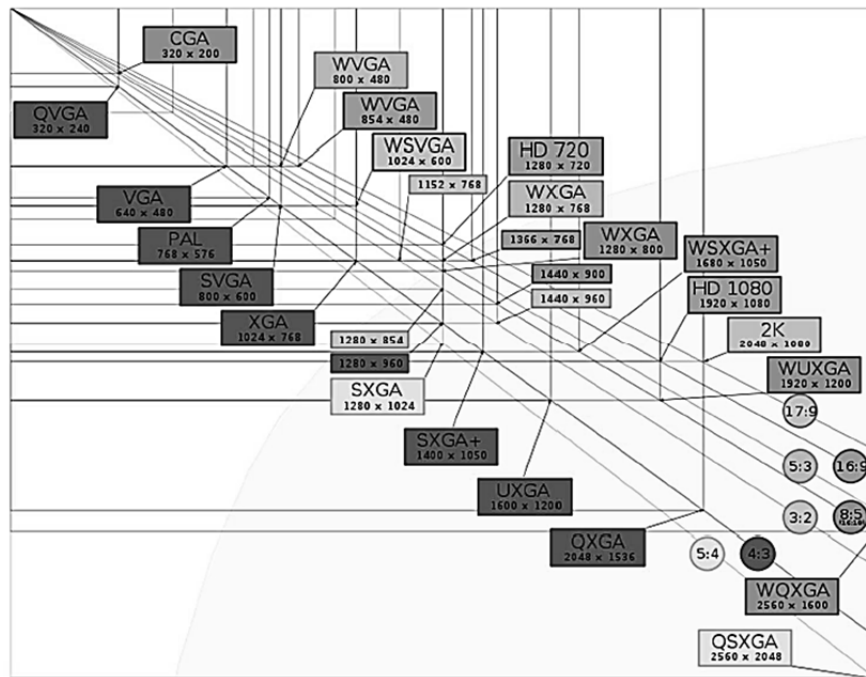


Figura 4-3. Comparación de resoluciones entre los diferentes formatos de vídeo.

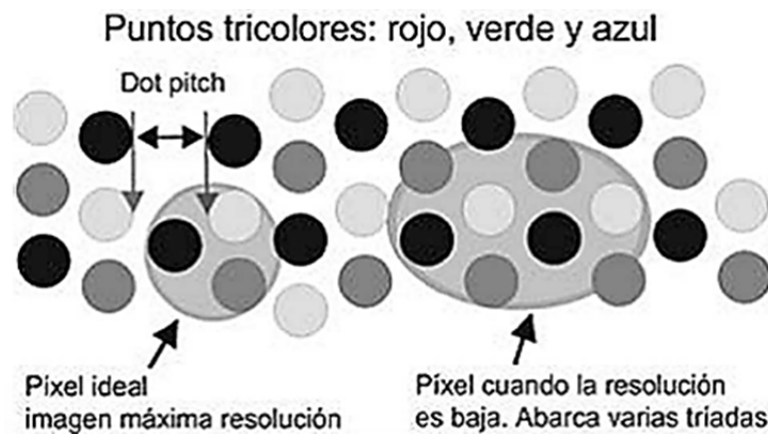


Figura 4-4. Tamaño del punto.

Persistencia de la pantalla

La persistencia de la pantalla indica la capacidad de la pantalla para mantener la emisión luminosa desde que un punto de la misma es excitado. Este parámetro se obtiene midiendo el tiempo transcurrido desde que un punto de la pantalla es excitado hasta que la emisión luminosa en ese punto disminuye al 1% del valor máximo. La persistencia debe ser lo suficientemente elevada como para que no se produzca la sensación de parpadeo pero debe ser inferior a la frecuencia de cuadro con objeto de que un cuadro no persista en el siguiente.

Pureza del color

La pureza del color indica la capacidad de un monitor de mostrar un color con una tonalidad uniforme en toda pantalla.

Número de colores

Un monitor puede representar de forma simultánea en pantalla un número de colores limitado. Cada punto o píxel debe poder representar la mayor gama de colores posibles. Por ejemplo, con 8 bits de color se pueden representar 16,7 millones de colores.

Ángulo de visión

El ángulo de visión de un monitor es el ángulo total dentro del cual se obtiene una correcta visión de la imagen en el monitor. Interesan valores altos, próximos a 180°. Del ángulo de visión dependen los parámetros de brillo, contraste y color, que disminuyen a medida que aumenta el ángulo.

Tiempo de respuesta

El tiempo de respuesta es el que se requiere para pasar de un estado de brillo a otro y volver al inicial. Interesa que sea lo más bajo posible. Existen dos tipos de medidas: blanco a negro y gris a gris (ésta última es la más representativa en imágenes reales). Los valores más comunes oscilan entre 2 ms y 8 ms.

4.2 Conectores

Según el tipo de señal se escogerá un tipo de conector determinado.

a) Vídeo: en el que se usa con la señal analógica (señal compuesta de videofrecuencia, SCVF). Se usan conectores RCA identificados con el color amarillo. Como se muestra en la Figura 4-5 van acompañados de dos conectores RCA correspondientes al audio (Rojo-canal derecho y Blanco-canal izquierdo).



Figura 4-5. Conector RCA para SCVF. Hembra RCA (arriba) y macho RCA (abajo).

b) S-Vídeo

Formato de señal analógica que trabaja con luminancia y crominancia. Se usa el conector Mini-DIN de 4 pins representado en la Figura 4-6.



Figura 4-6. Conector para S-Vídeo: DIN 4 pins.

c) Vídeo en componentes

Formato de señal analógica que trabaja en componentes Y Pb Pr. Se usan tres conectores RCA: Verde (Y), azul (Pb/Cb) y rojo (Cb/Cr).



Figura 4-7. Conectores RCA para vídeo en componentes.

d) VGA (Video Graphics Array)

Formato de señal analógica con tres canales separados para los colores primarios y para las señales de sincronismo. El conector usado para este tipo de formato es el SUB-D15. En la Figura 4-8 se muestra la hembra de este tipo de conector.

Patilla	Función	Patilla	Función
1	Rojo	9	+5V (con fusible de 250 mA)
2	Verde	10	Conexión a tierra (ret. sinc.)
3	Azul	11	No conectado
4	No conectado	12	SDA (datos DDC)
5	Conexión a tierra	13	Sinc. horizontal
6	Retorno rojo (conex tierra)	14	Sinc. vertical
7	Retorno verde (conex tierra)	15	SCL (reloj DDC)
8	Retorno azul (conex a tierra)		



Figura 4-8. Conector SUB-D15 para VGA.

e) **HDMI** (*High-Definition Multimedia Interface*): Interfaz de audio y vídeo digital no comprimido estándar y adoptado globalmente por la industria para sistemas de HD. Cable único para audio, vídeo y datos. El conector es específico para este formato y consta de 19, dispuestos como muestra la Figura 4-10 y con funciones especificadas en la Tabla 4-3. Los datos se codifican en formato TMDS (*Transition Minimized Differential Signaling*), una tecnología de transmisión de datos en serie a alta velocidad.

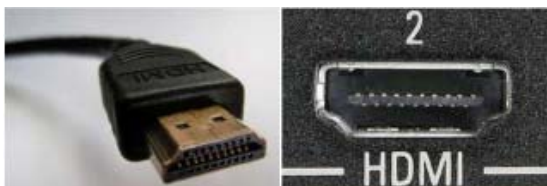


Figura 4-9. Conector para HDMI.

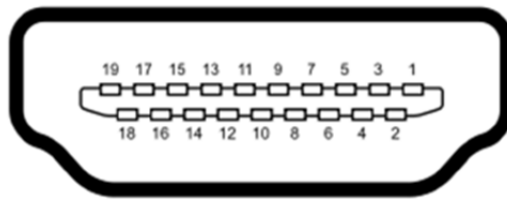


Figura 4-10. Pines del HDMI.

Tabla 4-1. Conexión del HDMI.

Pin 1	TMDS Data2+
Pin 2	TMDS Data2 Shield
Pin 3	TMDS Data2–
Pin 4	TMDS Data1+
Pin 5	TMDS Data1 Shield
Pin 6	TMDS Data1–
Pin 7	TMDS Data0+
Pin 8	TMDS Data0 Shield
Pin 9	TMDS Data0–
Pin 10	TMDS Clock+
Pin 11	TMDS Clock Shield
Pin 12	TMDS Clock–
Pin 13	CEC
Pin 14	Reserved (N.C. on device)
Pin 15	SCL
Pin 16	SDA
Pin 17	DDC/CEC Ground
Pin 18	+5 V Power (max 50 mA)
Pin 19	Hot Plug Detect

f) DVI (Digital Visual Interface).

Formato de vídeo digital que incluye también datos auxiliares. Los datos se codifican en formato TMDS como en HDMI. Por eso una fuente DVI puede conectarse a un monitor HDMI, o viceversa, por medio de un cable adecuado, pero el audio y las características del control remoto HDMI no estarán disponibles. Utiliza un conector de 24 pines, cuya hembra viene representada en la Figura 4-11.



Figura 4-11. Conector DVI-D usado solo para señales digitales.

Los conectores DVI se clasifican en función de las señales que admiten:

- DVI-D (sólo digital)

- DVI-A (sólo analógica)
- DVI-I (digital y analógica)

Se denomina DVI-DL a los conectores que admiten dos enlaces. El conector DVI tiene pines para:

- Transmitir señales digitales nativas.
- En los sistemas de doble enlace, hay pines adicionales para la segunda señal.
- También puede tener pines para transmitir las señales analógicas del estándar VGA.

En la Tabla 4-2 se especifica la función de cada pin del conector.

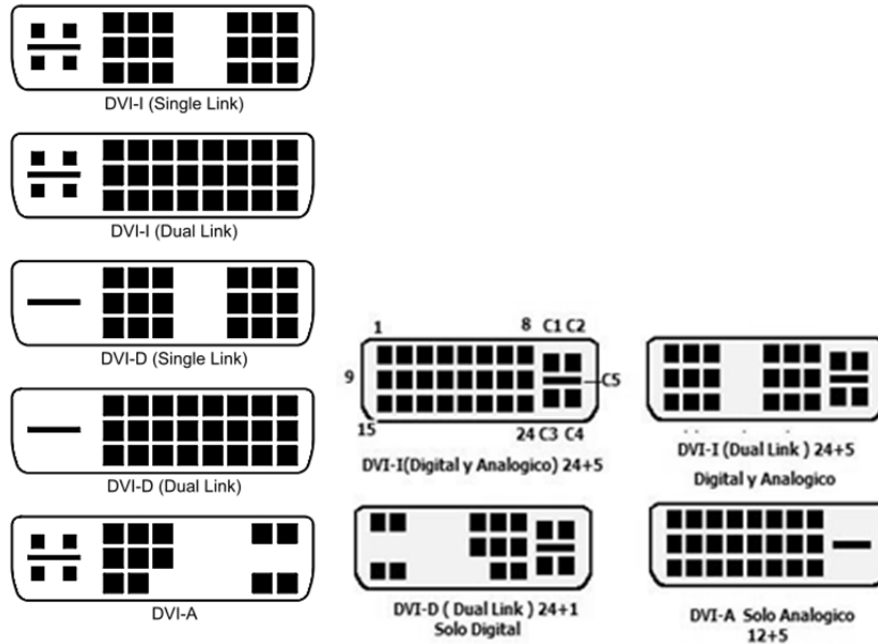


Figura 4-12. Pineado del DVI dependiendo del tipo de señal que se transmita.

Tabla 4-2. Señal transportada en cada pin del DVI.

Pin	Nombre	Función
1	Datos TMDS 2-	Rojo digital - (Link 1)
2	Datos TMDS 2+	Rojo digital + (Link 1)
3	Protección datos TMDS 2/4	
4	Datos TMDS 4-	Verde digital - (Enlace 2)
5	Datos TMDS 4+	Verde digital + (Enlace 2)
6	Reloj DDC	
7	Datos DDC	
8	Sincronización vertical analógica	
9	Datos TMDS 1-	Verde digital - (Enlace 1)
10	Datos TMDS 1+	Verde digital + (Enlace 1)
11	Protección datos TMDS 1/3	
12	Datos TMDS 3-	Azul digital - (Enlace 2)
13	Datos TMDS 3+	Azul digital + (Enlace 2)
14	+5V	Energía para el monitor en espera
15	Tierra	Retorno para pin 14 y sincronización analógica
16	Detección Hot Plug	
17	Datos TMDS 0-	Azul digital - (Enlace 1) y sincronización digital
18	Datos TMDS 0+	Azul digital + (Enlace 1) y sincronización digital
19	Protección datos TMDS 0/5	
20	Datos TMDS 5-	Rojo digital - (Enlace 2)
21	Datos TMDS 5+	Rojo digital + (Enlace 2)
22	Protección reloj TMDS	
23	Reloj TMDS +	Reloj digital + (Enlaces 1 y 2)
24	Reloj TMDS -	Reloj digital - (Enlaces 1 y 2)
C1	Rojo analógico	
C2	Verde analógico	
C3	Azul analógico	
C4	Sincronización horizontal analógica	
C5	Tierra (analógico)	Retorno para señales de Rojo, Verde y Azul
C1	señal analógica Rojo	
C2	señal analógica Verde	
C3	señal analógica Azul	
C4	señal Sincronismo horizontal	
C5	masa señal Sincronismo(RGB)	

4.3 Tecnologías usadas en Monitores

4.3.1 Monitores CRT (Tubo de rayos catódicos)

El **tubo de rayos catódicos**, en la Figura 4-13, (CRT, del inglés *Cathode Ray Tube*) es una tecnología que permite visualizar imágenes mediante un haz de rayos catódicos constante dirigido contra una pantalla de vidrio recubierta de fósforo y plomo. El fósforo permite reproducir la imagen proveniente del haz de rayos catódicos, mientras que el plomo bloquea los rayos X para proteger al usuario de sus radiaciones. Fue desarrollado por William Crookes en 1875, aunque no se utilizó hasta 1940 con la aparición de los primeros televisores. Se emplea principalmente en monitores, televisores y osciloscopios, aunque en la actualidad han sido sustituidos por tecnologías como el plasma, el LCD, el LED o el DLP.

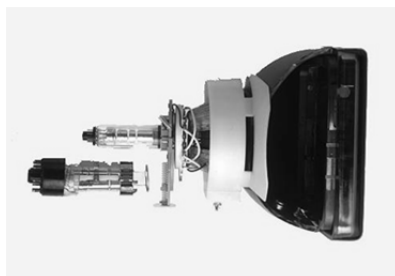


Figura 4-13. Interior de una pantalla de televisión.

4.3.2 Monitores LCD

Las Pantallas de Cristal Líquido (LCD del inglés, Liquid Cristal Display), están altamente difundidas en la actualidad. Son muy útiles porque permiten mostrar información o datos de manera muy clara. La mayoría de los electrodomésticos y diversos equipos electrónicos traen uno o varios de ellos porque presentan la gran ventaja de consumir poco.

El principio de funcionamiento en el que se basa el LCD no es la emisión de luz sino el bloqueo de la misma. Los LCD se basan en cuatro principios:

- La luz se puede polarizar.
- Los cristales líquidos pueden transmitir y cambiar la luz polarizada.
- La estructura de un cristal líquido puede ser cambiada por corriente eléctrica.
- Hay sustancias transparentes que pueden conducir electricidad.

Existen diferentes tecnologías utilizadas por los monitores planos, siendo la del LCD una de las más utilizadas en monitores para ordenadores por su bajo consumo.

Funcionamiento del Cristal Líquido

Los cristales líquidos representan un estado intermedio de la materia entre cristalino y líquido. Las moléculas de LC (*Liquid Cristal*), representadas en la Figura 4-14, tienen forma de barra plana o disco y sus átomos se sitúan a lo largo del eje de la molécula.

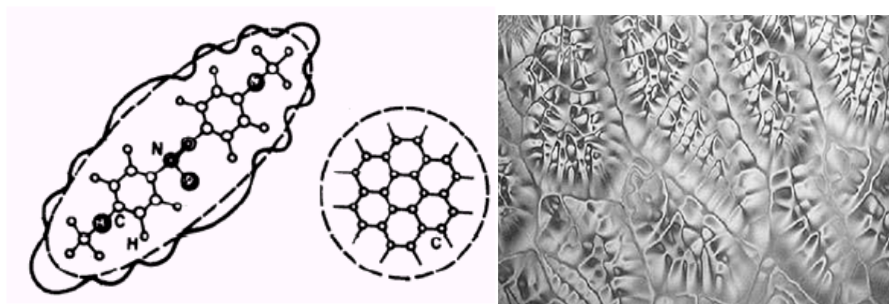


Figura 4-14. Estructura de una molécula LC.

El tipo de LC que se utiliza para hacer las pantallas son los nemáticos. Las moléculas en los cristales líquidos nemáticos tienen un orden definido. Una cualidad de los cristales líquidos es que se ven afectados por la corriente eléctrica. Un tipo particular de cristal líquido es el Twisted Nematics y se caracteriza porque cuando se le aplica una corriente eléctrica las moléculas se “desenredan” (untwist) varios grados, dependiendo del voltaje aplicado. Este tipo de cristal líquido es usado en las pantallas, porque su reacción a la corriente es moldeable y permite controlar el paso de la luz.

Las moléculas, al igual que en una fase líquida normal se encuentran esencialmente desordenadas en cuanto a las posiciones de sus centros de masa, pero en alguno de los ejes principales de las moléculas (por ejemplo, su eje longitudinal) las moléculas se encuentran en promedio orientadas en una determinada dirección, llamada director, tal y como se representa en la Figura 4-15.

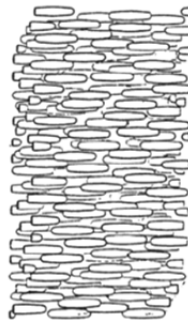


Figura 4-15. Fase nemática en la que hay orientación de las moléculas pero desorden en la ubicación de sus centros.

Se podría comparar este arreglo molecular con el de las cerillas dentro de una caja: aunque pueden moverse, la presencia de las demás impone restricciones a estos movimientos y tienden a mantenerse paralelas entre sí.

La peculiaridad de este tipo de materiales es que la existencia de una dirección preferida afecta al comportamiento de los rayos luminosos en el material cambiando su intensidad, color y dirección de propagación. Por esta razón a esta dirección especial se le llama el eje óptico del material y es la causa de muchos fenómenos ópticos importantes. Es necesario subrayar aquí que la existencia del eje óptico es un fenómeno colectivo que se da en forma espontánea en los cristales líquidos.

Por otra parte, es bien conocido que los cristales sólidos con ejes ópticos se usan en múltiples aplicaciones tecnológicas en las que es necesario dirigir los rayos luminosos de forma fija y precisa. Una aplicación muy conocida es la de los polaroides o polarizadores, que sólo transmiten algunos de los rayos luminosos que inciden sobre ellos y que se usan ampliamente en fotografía o para fabricar lentes que protejan los ojos de la luz. Para entender mejor la función de un polarizador recordemos que cada onda electromagnética emitida consta de un campo eléctrico E y un campo magnético H que oscilan transversalmente a la dirección de propagación de la onda, como se muestra en la Figura 4-16. Aunque en cada onda el campo eléctrico está orientado de modo preciso, en un rayo de luz compuesto por muchas ondas emitidas por diferentes sistemas de cargas en diferentes instantes de tiempo, no hay, en general, una dirección de polarización bien determinada del campo eléctrico total. Cuando esto ocurre se dice que la luz es natural o no polarizada.

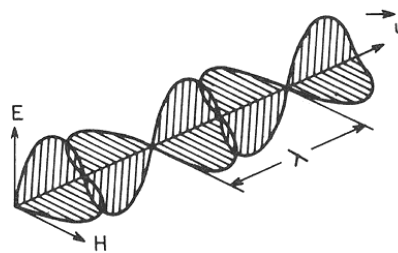


Figura 4-16. Onda electromagnética propagándose en la dirección de u . λ es la longitud de onda.

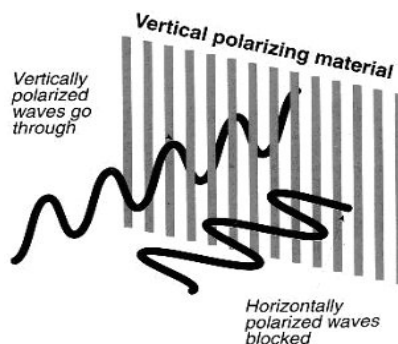


Figura 4-17. Polarización vertical.

Los polarizadores permiten, precisamente, convertir la luz natural en polarizada al eliminar del rayo de luz todas aquellas ondas cuya polarización no sea la requerida. Ya que los cristales líquidos poseen eje óptico no es de extrañar que también exhiban muchos de estos fenómenos, los cuales son totalmente insólitos en los líquidos ordinarios. Sin embargo, a diferencia de los cristales sólidos los cristales líquidos presentan una ventaja enorme: debido a la relativa debilidad de las fuerzas de interacción molecular, la orientación del eje óptico se puede variar con mayor facilidad. Existen diversos métodos para manejar el eje óptico de un cristal líquido, todos ellos basados en la enorme capacidad de estos sistemas para responder a la acción de campos externos.

El primer fenómeno en el que se observó que el eje óptico cambia de dirección bajo la acción de un campo eléctrico es el ahora llamado **efecto Freederiks** (1943), descubierto por el físico ruso del mismo nombre. Este fenómeno se observa con mayor facilidad en un nemático y, como se verá, constituye la base física del funcionamiento de los LCD. Como se representa en la Figura 4-18, si un campo eléctrico externo actúa sobre un nemático, interacciona con los dipolos de sus moléculas y las hace girar de manera tal que su eje largo tiende a colocarse paralelamente al campo aplicado o en dirección perpendicular a él, dependiendo de que el momento dipolar de la molécula esté dirigido en la dirección de su eje largo o corto, respectivamente. Pero ya hemos visto que en un nemático todas las moléculas mantienen una misma orientación debido a las características de su interacción mutua; así que si una molécula de nemático se reorienta bajo la acción de un campo externo, las demás se reorientan también en la misma dirección. Por lo tanto, basta con aplicar un campo eléctrico pequeño para producir una reorientación de todo el eje óptico del nemático.

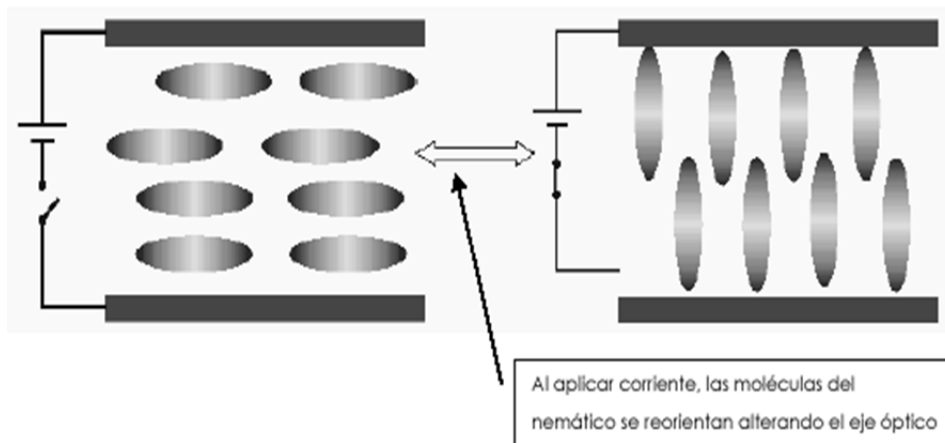


Figura 4-18. Reorientación de las moléculas del cristal líquido nemático.

Un indicador electro-óptico consta básicamente de una celda diminuta con los siguientes elementos: una capa muy delgada (0.01 - 0.1 mm) de cristal líquido, por lo general un nemático, colocado entre dos placas de vidrio cada una de las cuales está unida a un polarizador. Estas placas se pulen de manera tal que la cohesión entre las moléculas y las placas sólidas sea máxima, lo cual se logra produciendo surcos en las placas de aproximadamente las dimensiones de una molécula de nemático. Así se consigue que las moléculas que están en contacto con las placas se alineen con la dirección de pulimento y, por las características de las interacciones moleculares, todas las moléculas en el espesor de la capa se alinean exactamente en la misma dirección. Si bajo estas condiciones una placa se gira noventa grados respecto a la otra, debido a que a grandes distancias de las placas el influjo del vidrio sobre las moléculas de nemático es menor que cerca de las superficies sólidas, las moléculas, y por lo tanto el eje óptico del cristal líquido, se reorientan adquiriendo la configuración torcida como se muestra en la Figura 4-19.

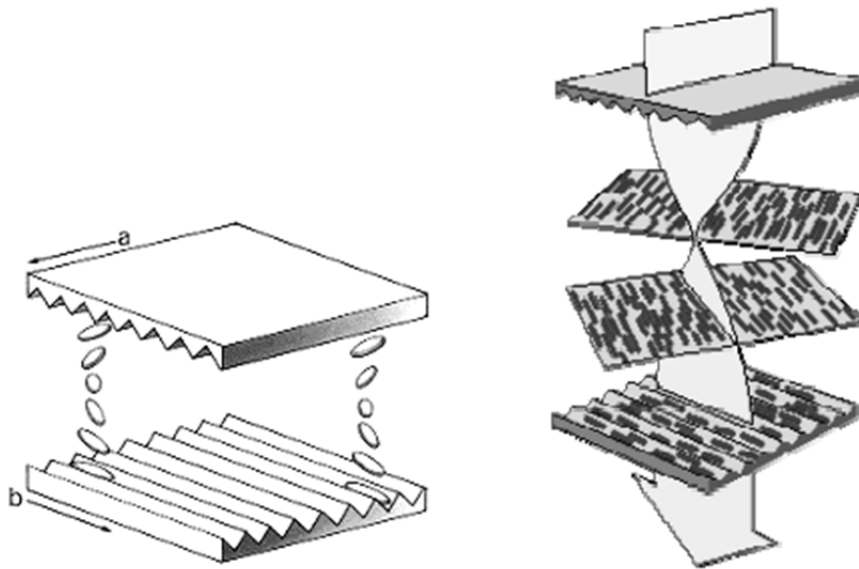


Figura 4-19. Geometría torcida producida al girar las placas de vidrio 90°.

Si entonces se ilumina la celda con luz natural, ésta se polariza y al penetrar en la celda su polarización sigue la configuración torcida del eje óptico, de modo que al salir por la placa inferior también habrá girado noventa grados. Pero como el eje del polarizador inferior tiene precisamente esa dirección, la luz atravesará el segundo polarizador. Entonces, si debajo del polarizador inferior se coloca un espejo, como se observa en la Figura 4-20, la luz podrá reflejarse nuevamente al interior de la celda, atravesarla en sentido inverso y salir de ella hasta llegar a los ojos de un observador que verá la celda perfectamente clara y transparente.

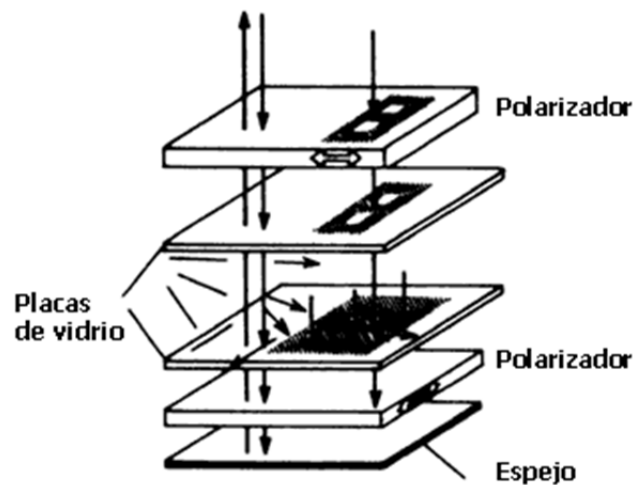


Figura 4-20. Elementos de una celda de un indicador electro-óptico.

Para poder formar cualquier cifra, letra o palabra en el indicador, es preciso colocar dos electrodos entre las placas de vidrio de manera que, al conectarlos a una pila, se produzca un campo eléctrico en el nemático. De hecho, el electrodo inferior se construye de una sola pieza, mientras que el superior consta de muchos electrodos más pequeños. Así, por ejemplo, para formar cualquier cifra entre cero y nueve sólo se requieren siete segmentos, cada uno de los cuales puede conectarse o desconectarse a un circuito eléctrico de acuerdo a las señales que produzca un generador. De esta manera es posible producir un campo eléctrico únicamente sobre ciertas regiones de la celda. En la Figura 4-20 se ilustra la situación en que ha sido cerrado el circuito de manera que los segmentos del electrodo superior que conforman el número ocho se activen. De acuerdo con el efecto Freederiks, el campo eléctrico en esta región de la celda tiende a alinear las moléculas destruyendo la configuración torcida que prevalecía en ausencia de campo. Entonces, la luz polarizada que sale del primer

polarizador se transmite, por debajo de los segmentos de electrodo asociados a la cifra ocho, por toda la capa de nemático sin cambiar su polarización. Pero al llegar al segundo polarizador, cuya dirección de polarización es perpendicular a la polarización de la luz, no podrá atravesarlo y entonces o se absorbe en el líquido o se desvía en cualquiera otra dirección. Como esta luz no llega al espejo, tampoco alcanzará al observador y éste verá oscuros los segmentos de la cifra ocho. El resultado final será que el observador verá un ocho oscuro sobre un fondo claro tal y como se ve en la Figura 4-21.



Figura 4-21. Imagen típica de un indicador digital electro-óptico.

Es importante señalar que los procesos comprendidos en los dispositivos que se acaban de describir, ocurren rápidamente, por lo común en una diezmilésima o cienmilésima de segundo; así, en una fracción pequeña de tiempo se puede mostrar en la pantalla toda una sucesión de cifras o símbolos diversos.

Por lo tanto, la elaboración de este tipo de pantalla se basa en la unión de una serie de capas que realizan trabajos independientes. El paso o no de la luz a través de ellas da como resultado la imagen a visualizar.

En la Figura 4-22 se detallan las capas que tiene una pantalla LCD.

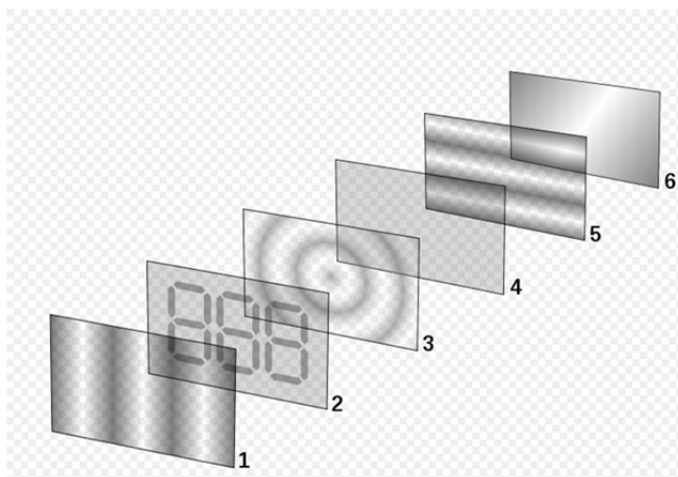


Figura 4-22. Capas de una pantalla LCD.

1-Capa de filtro vertical para polarizar la luz cuando entra.

2-Sustrato de vidrio con electrodos de Oxido de Indio y Estaño (ITO *electrodes*). Las formas de estos electrodos son las que crean las franjas oscuras que aparecen cuando se enciende el LCD. Las líneas verticales están grabadas en la superficie lo que supone que los cristales líquidos estén en línea con la luz polarizada.

3-Cristal líquido nemático en forma circular (trenzado).

4-Sustrato de cristal con una película de electrodo común (ITO) con crestas horizontales para alinearse con el filtro horizontal.

5-Película de filtro horizontal para bloquear/permitir el paso de la luz.

6-Superficie reflectante para hacer que la luz vuelva al visor.

Dependiendo de la posición relativa de la fuente de luz (tanto natural como artificial) que pase a través del cristal líquido, se definen tres tipos claramente diferenciados:

-Transmisivo, empleada en los televisores LCD.

- Reflexiva, utilizada en las calculadoras y relojes.
- .Proyectada, generada por proyectores LCD.

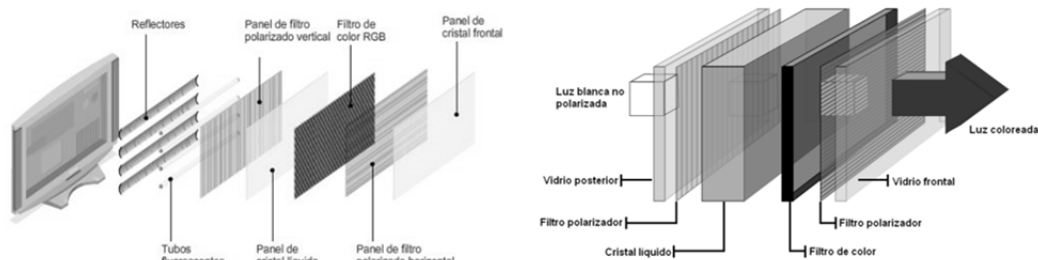


Figura 4-23. Capas de un monitor LCD: modo transmisivo.

Direccionamiento

Para iluminar un píxel de la pantalla, los monitores LCD usan métodos de direccionamiento basados en matrices activas y pasivas que se detallan a continuación:

a) Monitores LCD de matriz pasiva

Están formados por dos filtros polarizadores con filas de electrodos transparentes situados perpendicularmente por encima y por debajo del LCD. Aplicando una corriente eléctrica a los filtros se consigue que la luz pase o no dependiendo de que lo permita o no el segundo filtro. En la Figura 4-24 se muestra la estructura de las matrices pasivas.

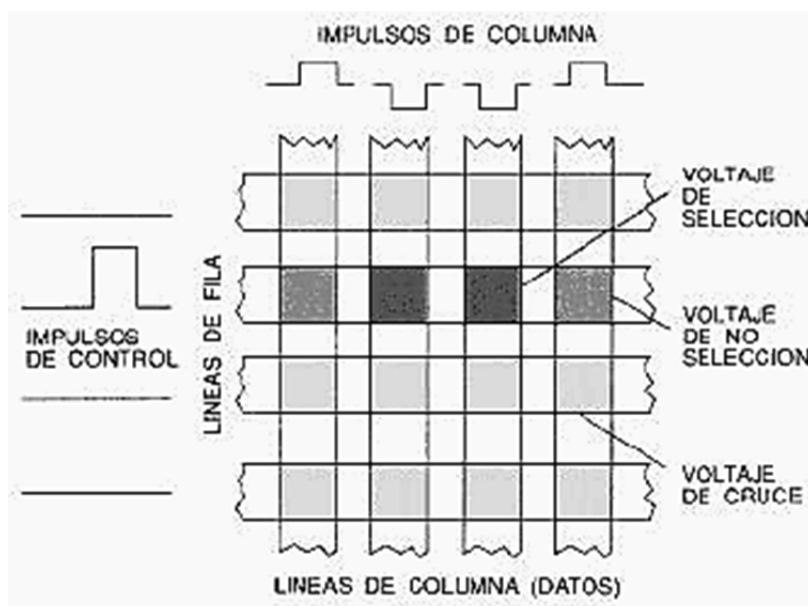


Figura 4-24. Estructura de una matriz pasiva de LCD.

Si se intercalan tres filtros adicionales de colores básicos (rojo, verde, azul), se obtienen pantallas que reproducen imágenes en color. Ésta es la base de las **pantallas DSTN**, o de matriz pasiva, que se emplearon en ordenadores portátiles y otros dispositivos móviles, porque tenían ventajas. Al carecer de tubo, las pantallas DSTN eran completamente planas, mucho más delgadas, consumían menos energía y producían menos emisiones radiactivas. El problema que presentan es una imagen fantasma o rastro de lo proyectado.

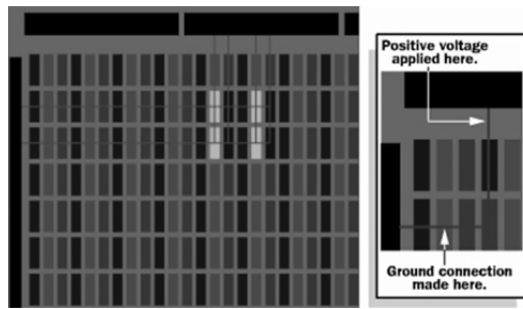


Figura 4-25. Pantalla de matriz pasiva con filtros de colores básicos.

b) Monitores de matriz activa

TFT-LCD (*Thin Film Transistor-Liquid Crystal Display*, Transistor de Película Fina - Pantalla de Cristal Líquido) es una variante de pantalla de cristal líquido (LCD) que usa tecnología de transistor de película delgada (TFT) para mejorar su calidad de imagen. Las LCD de TFT son un tipo de LCD de matriz activa, aunque esto es generalmente sinónimo de LCD. Son usados en televisores, visualizadores de pantalla plana y proyectores. En informática, los monitores de TFT han desplazado la tecnología de CRT, y están comúnmente disponibles en tamaños de 12 a 30 pulgadas.

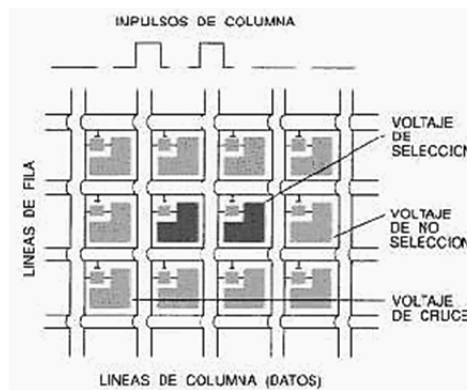


Figura 4-26. Estructura de una pantalla de matriz activa. Cada píxel (transistor) puede ser activado individualmente sin afectar a píxeles contiguos.

Las pantallas de cristal líquido normales, como las de las calculadoras, presentan elementos de imagen excitados en forma directa, se puede aplicar una tensión a través de un segmento sin que interfiera con otros segmentos de la pantalla. Esto no es posible en pantallas grandes con un gran número de píxeles, puesto que se requerirían millones de conexiones, conexiones en la parte superior e inferior para cada uno de los tres colores (rojo, verde y azul) de cada píxel. Para evitar esto, los píxeles son direccionados en filas y columnas, lo que reduce el número de conexiones de millones a miles como se puede observar en la Figura 4-27.

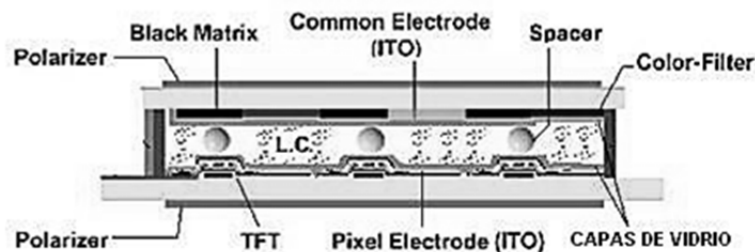


Figura 4-27. Sección de una pantalla LCD TFT. En una pantalla de color con una resolución de 1024x768 hay un total de $1024 \times 3 \times 768 = 2.359.296$ transistores.

Si todos los píxeles de una fila son excitados mediante una tensión positiva y todos los píxeles de una columna son excitados con una tensión negativa, entonces el píxel que se encuentra en la intersección tiene el voltaje aplicado más elevado y es conmutado. El inconveniente de esta solución es que todos los píxeles de la misma columna reciben una fracción de la tensión aplicada, como ocurre con todos los píxeles de la misma fila, así a pesar de que no sean conmutados completamente, tienden a oscurecerse. La solución al problema es proporcionar a cada píxel su propio transistor conmutador, esto permite controlar a cada píxel por separado. La baja corriente de fuga del transistor implica que la tensión aplicada al píxel no se pierde durante las actualizaciones de refresco de la imagen en la pantalla. Cada píxel es un pequeño condensador con una capa transparente de óxido de indio y estaño en el frontal, una capa transparente en la parte posterior, y entre medio una capa aislante de cristal líquido. Requieren una iluminación mayor ya que el TFT absorbe luz. En comparación con los monitores CRT tienen menor contraste, luminosidad y menor velocidad de actualización. Pero añaden la ventaja del tamaño reducido y peso ligero.

Distribución de los píxeles

1. Tiras: los colores están agrupados por columnas, son de fabricación y diseño simples. Presentan mezcla pobre de color y poca resolución.

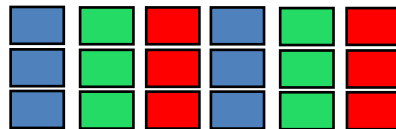


Figura 4-28. Distribución en tiras de los píxeles.

2. Mosaico: los colores están agrupados en diagonal. Cada subpíxel de un color está rodeado por otros colores diferentes. Presentan mezcla buena de color y buena resolución. Fabricación más compleja.



Figura 4-29. Distribución en mosaico de los píxeles.

3. Delta: cada fila se desplaza medio subpíxel respecto de la anterior. Cada subpíxel se encuentra en el centro de un hexágono cuyos vértices son colores diferentes al suyo. Presenta mejor característica de color y resolución. Es el más difícil de fabricar.

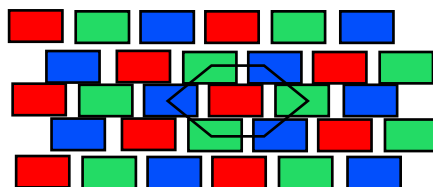


Figura 4-30. Distribución en Delta de los píxeles.

4.3.2.1 Proyectores LCD

El sistema de pantalla de cristal líquido es el más simple, por tanto uno de los más comunes y accesibles para el uso doméstico. En esta tecnología, la luz se divide en tres y se hace pasar a través de tres paneles de cristal líquido, uno para cada color fundamental (rojo, verde y azul); finalmente las imágenes se recomponen en una, constituida por píxeles, y son proyectadas sobre la pantalla mediante un objetivo.

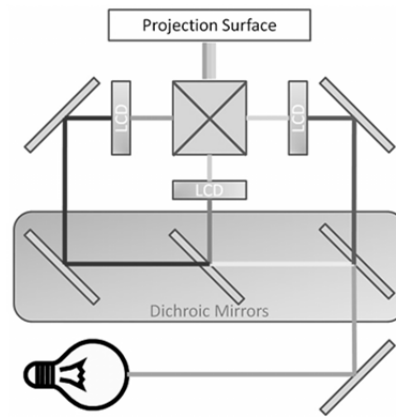


Figura 4-31. Esquema de proyectores LCD.

- Ventajas: es más eficiente que otros sistemas (imágenes más brillantes) y produce colores muy saturados.
- Inconvenientes: es visible un efecto de pixelación (aunque los avances más recientes en esta tecnología lo han minimizado), es probable la aparición de píxeles muertos y la vida de la lámpara es de aproximadamente 2000 horas.

4.3.3 Monitores LED

Las siglas LED corresponden a *Light Emitting Diode* y, en pocas palabras, se trata de conseguir luz a partir de componentes electrónicos en vez de los métodos tradicionales como la fluorescencia o la incandescencia halógena.

Una pantalla LED es un dispositivo compuesto de paneles o módulos de LED (diodos emisores de luz) debidamente compuestos por LEDs RGB. Se utilizan los LEDs disponiéndolos en forma de matriz utilizando diodos de distintos colores RGB para formar el píxel, de tal forma que se pueden obtener pantallas de LED de diversos tipos como lo serían: tipo indicadores, informativas, publicitarias y de alta resolución de vídeo a todo color. Un monitor LED es en realidad un monitor LCD al que se añade la tecnología LED. Aunque se habla de «monitores LED» y «monitores LCD», en realidad ambos son LCD. La iluminación LED sustituye a las lámparas fluorescentes.

Esta fuente de luz presenta más ventajas que la tradicional CCFL (*Cold Cathode Fluorescent Light*) usada en los LCD - TFT, exceptuando el factor coste, y caracteriza de forma determinante a las pantallas que integran esta fuente de luz como retroiluminación. El inconveniente de la iluminación por CCFL es que los colores de la pantalla que aprecia el ojo humano son algo artificiales y los tonos negros son grisáceos debido a la falta de precisión que esta tecnología obtiene a la hora de oscurecer las distintas zonas de visionado.

La tecnología LED permite retroiluminar porciones pequeñas de la pantalla en lugar de franjas enteras, con lo que se consigue una iluminación mayor y se mejora el contraste real. Se pueden «apagar» los diodos en las zonas donde no sean necesarios y generar «verdaderas» zonas negras. Se reduce el grosor de los monitores, se disminuye en un 50% el consumo eléctrico y se crea una mayor uniformidad del brillo y de intensidad.

4.3.4 Monitores plasma o PDP

Una **pantalla de plasma** (PDP: *plasma display panel*) es un dispositivo de pantalla plana habitualmente usado en televisores que consta de una matriz de celdas, cada una de las cuales se componen a su vez de tres sub-celdas «sub-píxeles», correspondientes a los colores rojo, verde y azul, situadas entre dos paneles de cristal que contienen una mezcla de gases nobles (neón y xenón). El gas en las celdas se convierte eléctricamente en plasma, el cual provoca que una sustancia fosforescente (que no es fósforo) pueda emitir luz. En la Figura 4-32 se puede ver la estructura de este tipo de monitores.

La tecnología de plasma fue concebida en 1964 en la Universidad de Illinois (USA). Desde entonces ha evolucionado de una manera espectacular. De unas pantallas de pequeño tamaño y una calidad más que deficiente, hemos pasado a grandes displays con alta calidad de color, brillo y contraste.

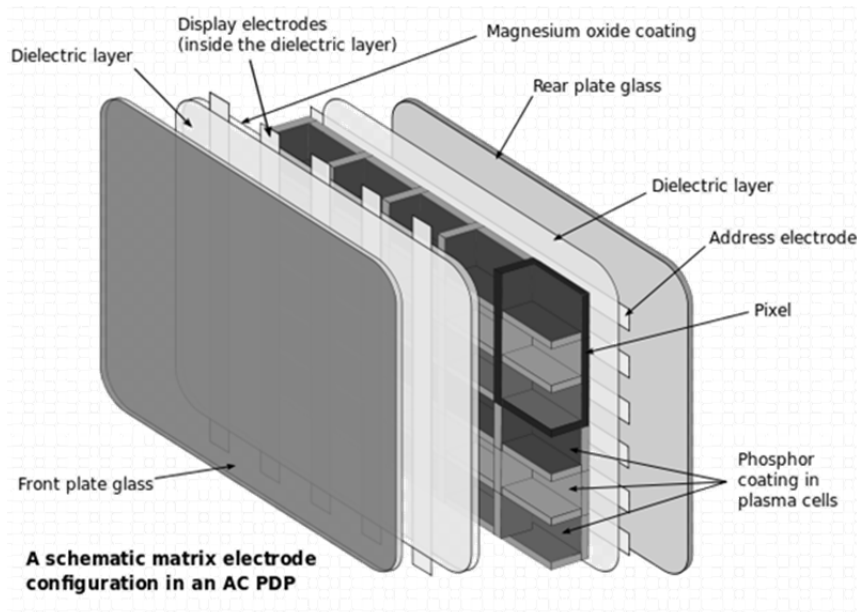


Figura 4-32. Composición de una pantalla de plasma.

El principio de funcionamiento se basa en «encender» mas o menos las pequeñas luces de colores (primarios R, G y B) para posteriormente por mezcla aditiva conformar el color del correspondiente píxel de la imagen.

Los gases xenón y neón en un televisor de plasma están contenidos en cientos de miles de celdas diminutas entre dos pantallas de cristal. Los electrodos también se encuentran «emparedados» entre los dos cristales, en la parte frontal y posterior de las celdas. Ciertos electrodos se ubican detrás de las celdas, a lo largo del panel de cristal trasero, y los otros electrodos, que están rodeados por materiales aislantes dieléctricos y cubiertos por una capa protectora de óxido de magnesio, están ubicados en frente de la celda, a lo largo del panel de cristal frontal. El circuito carga los electrodos que se cruzan creando diferencia de voltaje entre la parte trasera y la frontal, y provocan que el gas se ionice y forme el plasma. Posteriormente, los iones del gas corren hacia los electrodos, donde colisionan emitiendo fotones.

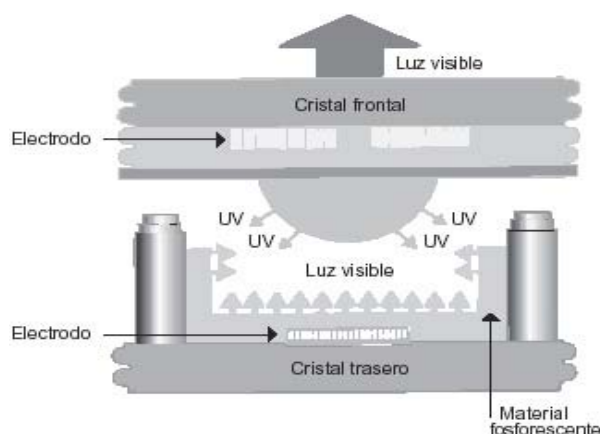


Figura 4-33. Principio de funcionamiento del plasma. Cada uno de los subpíxeles está formado por una pequeña celda estanca que contiene un gas inerte (neón o xenón). Al aplicar una diferencia de potencial entre los electrodos de la celda, dicho gas pasa al estado de plasma. El gas así cargado emite radiación ultravioleta (UV) que golpea y excita el material fosforescente que recubre el interior de la celda. Cuando el material fosforescente regresa a su estado energético natural, emite luz visible.

Las pantallas de plasma son brillantes (1000 lux o más por módulo), tienen una amplia gama de colores y pueden fabricarse en tamaños bastante grandes, hasta 262 cm de diagonal. Tienen una luminancia muy baja a nivel de negros, creando un negro que resulta más deseable para ver películas. Esta pantalla sólo tiene cerca de 6 cm de grosor y su tamaño total (incluyendo la electrónica) es menor de 10 cm. Los plasmas usan tanta energía por metro cuadrado como los televisores CRT. El consumo eléctrico puede variar en gran medida dependiendo de qué se esté viendo en él. Las escenas brillantes (como un partido de fútbol) necesitarán una mayor energía que las escenas oscuras (como una escena nocturna de una película). Las medidas nominales indican 400 vatios para una pantalla de 50 pulgadas. Los modelos relativamente recientes consumen entre 220 y 310 vatios para televisores de 50 pulgadas cuando se está utilizando en modo cine. El tiempo de vida de la última generación de pantallas de plasma está estimado en unas 100.000 horas (o 30 años a 8 horas de uso por día) de tiempo real de visionado; sin embargo, se han producido televisores de plasma que han reducido el consumo de energía y han alargado la vida útil del televisor. En concreto, éste es el tiempo de vida medio estimado para la pantalla, el momento en el que la imagen se ha degradado hasta la mitad de su brillo original. Se puede seguir usando pero se considera el final de la vida funcional del aparato.

La principal ventaja de la tecnología del plasma es que pantallas muy grandes pueden ser fabricadas usando materiales extremadamente delgados. Ya que cada píxel es iluminado individualmente, la imagen es muy brillante y posee un gran ángulo de visión.

En las pantallas electrónicas basadas en fósforo (incluyendo televisiones de rayos catódicos y de plasma), una exposición prolongada de una imagen estática puede provocar que los objetos que se muestren en ella queden marcados en la pantalla durante un tiempo. Esto es debido al hecho de que los compuestos fosforescentes que emiten la luz pierden su luminosidad con el uso. Como resultado, cuando ciertas áreas de la pantalla son usadas más frecuentemente que otras, a lo largo del tiempo las áreas de baja luminosidad se vuelven visibles a simple vista; esto se conoce como **pantalla quemada**, cuyo efecto sobre una pantalla se puede apreciar en la Figura 4-34. Un síntoma muy común es que la calidad de la imagen disminuye gradualmente conforme a las variaciones de luminosidad que tienen lugar a lo largo del tiempo, resultando una imagen con aspecto «emborronado».



Figura 4-34. Efecto pantalla quemada.

Las pantallas LCD, por el contrario, solían sufrir el denominado «efecto fantasma», algo desconocido en las pantallas CRT y plasma.

Como ventajas de las pantallas de plasma frente a las LCD se pueden destacar: mayor ángulo de visión, ausencia de tiempo de respuesta, lo que evita el efecto «estela» o «efecto fantasma» que se produce en ciertos LCD debido a altos tiempos de refresco (mayores a 12 ms), no contiene mercurio, a diferencia de las pantallas LCD, colores más suaves al ojo humano, un mayor número de colores y más reales y un contraste altísimo.

Como contraposición de las pantallas de plasma frente a las de LCD destaca que el coste de fabricación de los monitores de plasma es superior al de las pantallas LCD, este coste de fabricación no afecta tanto al PVP como al margen de ganancia de las tiendas, de ahí que muchas veces las grandes superficies no suelen trabajar con ellas, en beneficio de los de LCD. Además el consumo eléctrico de un televisor con pantalla de plasma grande

puede ser de hasta de un 30% más que un televisor LCD. No obstante, los nuevos plasmas tienen consumos muy razonables, del orden de los 140 W para un tamaño de 42".

4.3.5 PLED (Pixel Light Emitting Display)

Tecnología presentada en Abril de 2012 de la mano de LG y se basa en el mismo funcionamiento que las pantallas de plasma: cada píxel emite su propia luz. No requieren retroiluminación, se tiene control independiente del brillo de cada píxel consiguiéndose un mejor contraste. Las celdas son más pequeñas que un PDP y tienen mayor concentración de gas lo que supone menos energía para generar plasma, es decir menos consumo, menos desgaste y una mayor duración (+ 60%). Llevan tres procesadores que permiten intercalar imágenes (como los LED) y eliminar la sensación de ‘saltos’ del plasma convencional en imágenes con movimientos muy rápidos. En la Tabla 4-3 se establece una comparativa entre las tres tecnologías más similares entre si.

Tabla 4-3. Tabla comparativa entre las tecnologías más desarrolladas de monitores y televisión.

LED	PLED	Plasma
1. Puntos de luz dependen de la retroiluminación (200)	1. Puntos de Luz Independientes (6.220.800)	1. Puntos de luz Independientes (6.220.800)
2. Pocos Hz MCI, dejan Estela. Menos nítido	2. No deja Estela (0%). Mayor Nitidez	2. No Deja Estelas. Mayor Nitidez
3. Colores Artificiales, Menos Naturales	3. Mayor Reproducción de colores. Mas Naturales	3. Mayor reproducción de colores. Más Naturales
4. Refleja Según modelos	4. No hay reflejo	4. El Cristal Refleja
5. Menor Consumo Energético	5. Menor Consumo Energético	5. Consumo Energético
6. Imagen Estable	6. Imagen Estable	6. Parpadea la Imagen
7. Poco calentamiento	7. Poco calentamiento	7. La pantalla se calienta
8. 3D activos pierden la imagen con las gafas a 90°	8. 3D activos mantiene la imagen con las gafas a 90°	8. 3D activos mantiene la imagen con las gafas a 90°
9. El Negro es Azul/Gris	9. Negros 100%	9. Negros con reflejo
10. Más Caro	10. Más Económico	10. Más Económico
11. Diseño	11. Diseño	11. Diseño
12. Ligero	12. Ligero	12. Mayor Peso

4.3.6 Tecnología de visualización OLED

Un **diodo orgánico de emisión de luz**, también conocido como **OLED** o **SUPER LED** (acrónimo inglés de *organic light-emitting diode*), es un diodo que se basa en una capa electroluminiscente formada por una película de componentes orgánicos que reaccionan, a una determinada estimulación eléctrica, generando y emitiendo luz por sí mismos. Su estructura se puede ver en la Figura 4-35.

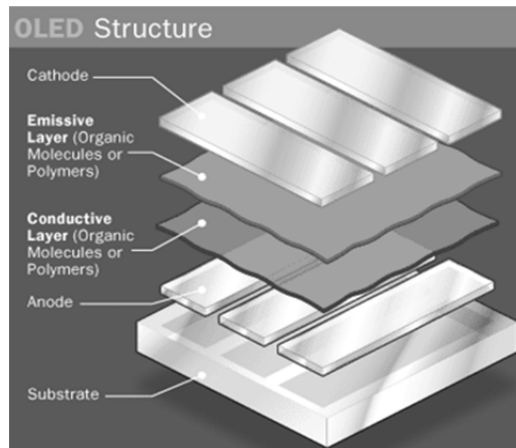


Figura 4-35. Estructura del OLED.

Existen muchas tecnologías OLED diferentes, tantas como la gran diversidad de estructuras (y materiales) que se han podido idear (e implementar) para contener y mantener la capa electroluminiscente, así como según el tipo de componentes orgánicos utilizados.

Un OLED está compuesto por dos finas capas orgánicas: una capa de emisión y una capa de conducción, que a la vez están comprendidas entre una fina película que hace de terminal ánodo y otra igual que hace de cátodo, como se aprecia en la Figura 4-36. En general estas capas están hechas de moléculas o polímeros que conducen la electricidad. Sus niveles de conductividad eléctrica se encuentran entre el nivel de un aislante y el de un conductor, y por ello se los llama semiconductores orgánicos.

La elección de los materiales orgánicos y la estructura de las capas determinan las características de funcionamiento del dispositivo: color emitido, tiempo de vida y eficiencia energética.

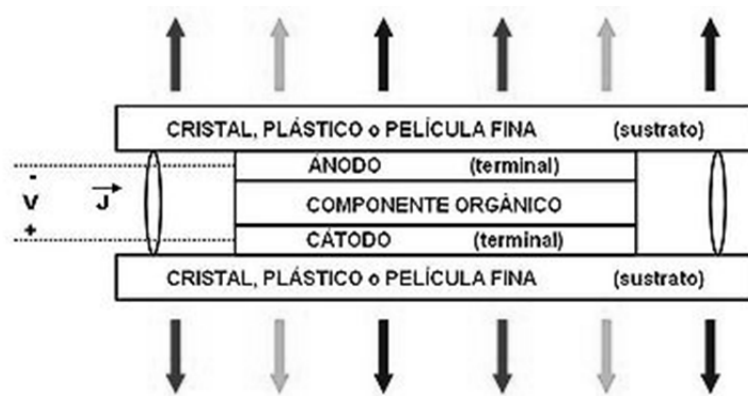


Figura 4-36. Capas de emisión y conducción del OLED.

El proceso de emisión de luz comienza cuando se aplica voltaje a través del OLED de manera que el ánodo sea positivo respecto del cátodo. Esto causa una corriente de electrones que fluye en sentido contrario de cátodo a ánodo. Así, el cátodo da electrones a la capa de emisión y el ánodo los sustrae de la capa de conducción.

Seguidamente, la capa de emisión comienza a cargarse negativamente (por exceso de electrones), mientras que la capa de conducción se carga con huecos (por carencia de electrones). Las fuerzas electrostáticas atraen a los electrones y a los huecos, los unos con los otros, y se recombinan (en el sentido inverso de la carga no habría recombinación y el dispositivo no funcionaría). Esto sucede más cerca de la capa de emisión, porque en los semiconductores orgánicos los huecos se mueven más que los electrones (no ocurre así en los semiconductores inorgánicos).

La recombinación es el fenómeno en el que un átomo atrapa un electrón. Dicho electrón pasa de una capa energética mayor a otra menor, liberándose una energía igual a la diferencia entre energías inicial y final, en forma de fotón.

La recombinación causa una emisión de radiación a una frecuencia que está en la región visible, y se observa un punto de luz de un color determinado. La suma de muchas de estas recombinaciones, que ocurren de forma simultánea, es lo que llamaríamos imagen. En la Figura 4-37 se expone de forma esquemática lo explicado anteriormente.

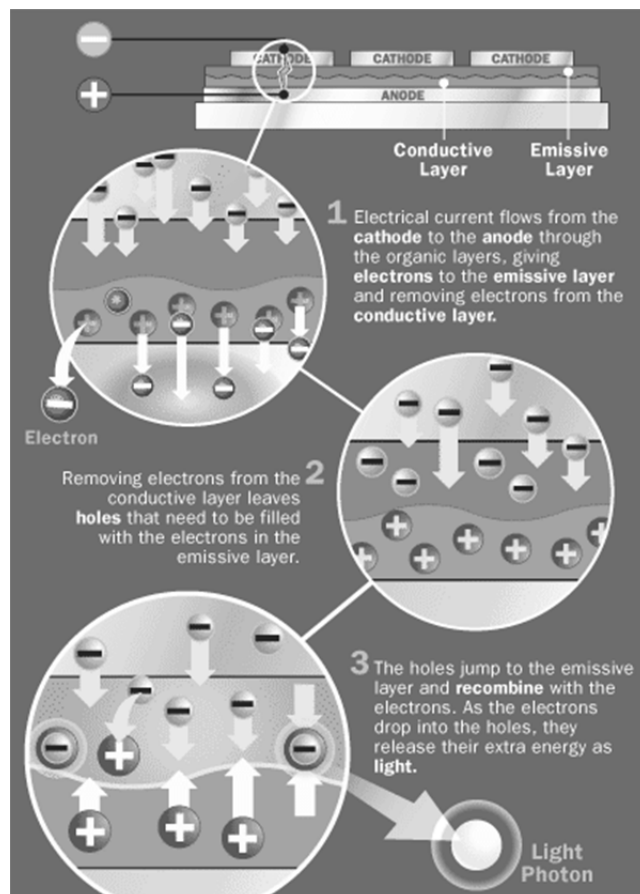


Figura 4-37. Principio de funcionamiento del OLED.

El color de la luz depende del tipo de molécula orgánica en la capa de emisión y la intensidad o brillo de la luz depende de la cantidad de corriente eléctrica aplicada.

Una característica especialmente interesante de los OLED es que son flexibles, Figura 4-38. Pueden doblarse sin dañarse, incluso en pleno funcionamiento. Su mayor desventaja es su limitado tiempo de vida.



Figura 4-38. La tecnología OLED permiten elasticidad en los elementos de visualización.

Actualmente se investiga para desarrollar nuevas versiones de OLED orgánico que no sólo emita luz, sino que también recojan la energía solar para producir electricidad y autoabastecerse.

Las aplicaciones principales del OLED son, entre otras, la de un periódico electrónico: una hoja flexible de plástico, en la que se pueden ver publicaciones obtenidas directamente desde Internet o una pantalla desarrollable para dispositivos móviles, videoconsolas y ordenadores portátiles.

Los OLED ofrecen muchas ventajas en comparación con los LCD, LED y pantallas de plasma.

Más delgados y flexibles: por una parte, las capas orgánicas de polímeros o moléculas de los OLED son más delgadas, luminosas y mucho más flexibles que las capas cristalinas de un LED o LCD. Por otra parte, en algunas tecnologías el sustrato de impresión de los OLED puede ser el plástico, que ofrece flexibilidad frente a la rigidez del cristal que da soporte a los LCD o pantallas de plasma.

Más económicos: en general, los elementos orgánicos y los sustratos de plástico serán mucho más económicos. También, los procesos de fabricación de OLED pueden utilizar conocidas tecnologías de impresión de tinta (en inglés, conocida como *inkjet*), hecho que disminuirá los costes de producción.

Brillo y contraste: los píxeles de los OLED emiten luz directamente. Por eso, respecto a los LCD posibilitan un rango más grande de colores y contraste.

Menos consumo: los OLED no necesitan la tecnología *backlight*, es decir, un elemento OLED apagado realmente no produce luz y no consume energía, a diferencia de los LCD que no pueden mostrar un verdadero “negro” y lo componen con luz consumiendo energía continuamente. Así, los OLED muestran imágenes con menos potencia de luz, y cuando son alimentados desde una batería pueden operar largamente con la misma carga.

Más escalabilidad y nuevas aplicaciones: la capacidad futura de poder escalar las pantallas a grandes dimensiones hasta ahora ya conseguidas por los LCD y, sobre todo, poder enrollar y doblar las pantallas en algunas de las tecnologías OLED que lo permiten, abre las puertas a todo un mundo de nuevas aplicaciones que están por llegar.

Mejor visión bajo ambientes iluminados: al emitir su propia luz, una pantalla OLED, puede ser mucho más visible bajo la luz del sol, que una LCD.



Figura 4-39. Comparativa entre TFT y OLED.

Desventajas y problemas actuales:

Tiempos de vida cortos: las capas OLED verdes y rojas tienen largos tiempos de vida, sin embargo la capa azul no es tan duradera, actualmente tienen una duración cercana a las 14.000 horas (8 horas diarias durante 5 años), este periodo de funcionamiento es mucho menor que el promedio de los LCD que dependiendo del modelo y del fabricante pueden llegar a las 60.000 horas.

Proceso de fabricación caro: actualmente la mayoría de tecnologías OLED están en proceso de investigación, y los procesos de fabricación (sobre todo inicialmente) son económicamente elevados, a no ser que se apueste por un diseño que se utilice en economías de escala.

Impacto medioambiental: los componentes orgánicos (moléculas y polímeros) se ha visto que son difíciles de reciclar (alto coste, complejas técnicas). Ello puede causar un impacto al medio ambiente muy negativo en el futuro.

4.3.7 Tecnología de visualización SED

El **Panel SED** es un tipo de panel visualizador para pantallas planas caracterizado por usar la tecnología de las pantallas de tubo tradicionales (CRT) para cada uno de los puntos (píxeles) mostrados en pantalla pero con el tamaño de los monitores LCD. En la Figura 4-40 se compara la tecnología CRT y SED. Cada píxel es un micro tubo de rayos catódicos. En principio, este tipo de paneles ofrece las ventajas de los tubos de imagen y los TFT, sin los defectos de ambos. De esta manera se consigue mejorar el contraste y el ángulo de visión sin aumentar el consumo. También permite ampliar las dimensiones de la pantalla con respecto a las pantallas de tecnología TFT o las de plasma. El proyecto ha sido desarrollado conjuntamente por Canon y Toshiba.

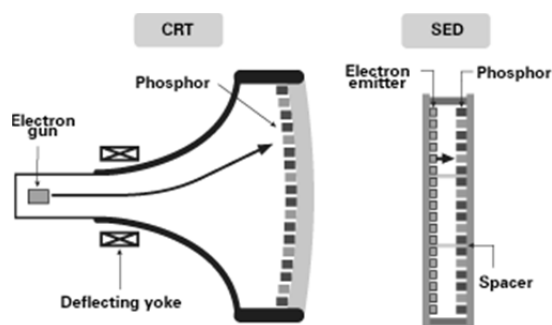


Figura 4-40. La diferencia entre SED y CRT es que mientras CRT utiliza un solo tubo para proyectar electrones, SED tiene tantos mini-proyectores como puntos de luz en pantalla.

La tecnología utilizada para estos paneles aún está en fase de desarrollo y no hay especificaciones de cómo va a funcionar exactamente. El principio radica en el de las televisiones normales de Tubo de Rayos Catódicos, (CRT). En los CRT, un haz de electrones es focalizado hacia cada píxel para iluminarlo según convenga. En la tecnología SED este procedimiento se "simplifica" porque se utiliza un haz de electrones individual para cada píxel, el cual iluminará posteriormente el fósforo encargado de producir la luz que hará brillar los colores primarios RGB de cada píxel. Gracias a esto no habrá que dirigir y focalizar un único rayo de electrones sobre una matriz de píxeles, sino que hay que montar millones de rayos de electrones sobre un panel SED. Uno para cada píxel. La estructura de un píxel SED se muestra en la Figura 4-41.

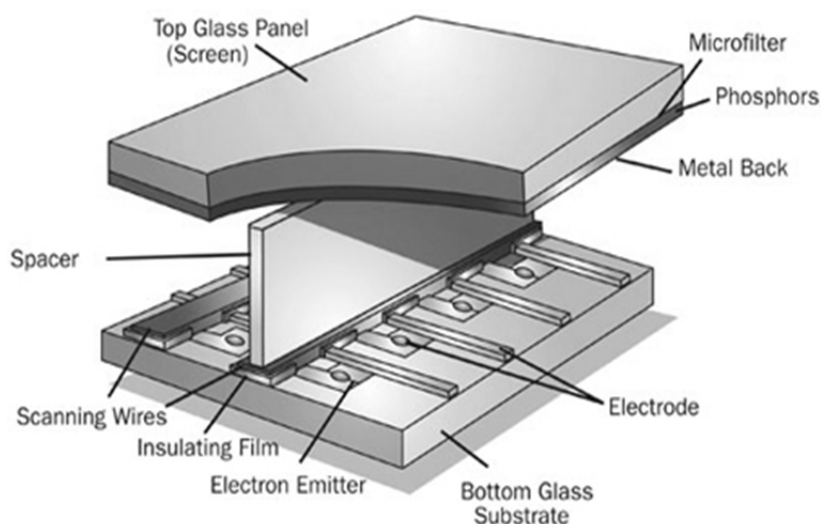


Figura 4-41. Estructura de un SED.

Así pues, cada píxel tendrá su propio cañón de electrones. Este cañón de electrones se forma a partir de dos electrodos muy pequeños (microscópicos) separados por unos nanómetros de distancia. A estos electrodos se les aplica una tensión de 16 voltios que "atravesará" los nanómetros que separa los electrodos. Fruto de este "salto"

entre electrodos, la corriente genera electrones que salen disparados hacia todas las direcciones. Para canalizar estos electrones hasta el fósforo que lo hará brillar se utiliza un campo eléctrico de 10 kV hacia la dirección donde esté el fósforo. De esta manera se genera un haz de electrones unidireccional que impacta sobre el fósforo perteneciente a un determinado píxel. Este proceso de iluminación se muestra en la Figura 4-42.

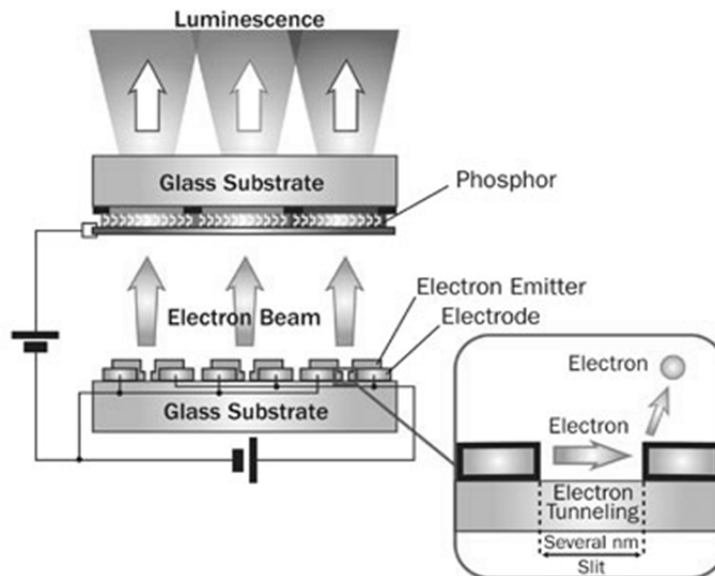


Figura 4-42. Proceso de iluminación de los píxeles en paneles SED.

Para formar una imagen entera se necesitan cientos de miles de píxeles (millones en alta definición). Por lo tanto actualmente se está estudiando la colocación de los cañones de electrones en un reducido espacio, sin que ello signifique una pérdida de funcionalidad por parte de dichos cañones o una pérdida de homogeneidad en la imagen.

Si combinamos las ventajas del LCD y del CRT obtenemos las ventajas de SED. Los tiempos de respuesta de los monitores SED son inferiores a los de los LCD, lo que permite mostrar imágenes más claras y con más definición. En la Figura 4-43 se compara la calidad de las tecnologías Plasma, SED y LED



Figura 4-43. Comparativa de imágenes en Plasma, SED y LCD.

El color, el contraste y la luminosidad serán iguales o mejores que los CRT. No habrá ningún problema para el ángulo de visión de la pantalla (no habrá cambios de color y brillo desde diferentes ángulos). Su consumo rebaja 2/3 la potencia utilizada en un CRT. 1/3 con respecto a la utilizada en los LCD. Tolerará temperaturas desde los -40° a los +85°. El proceso de fabricación es más sencillo que el de las pantallas LCD. No habrá problemas con el refresco de la imagen y su fluidez ya que utilizarán la misma velocidad de refresco que un televisor CRT normal. El color negro alcanzará mayor calidad.

4.3.8 Tecnología láser

La **pantalla láser** es una tecnología de visualización y proyección de vídeo basada en optoelectrónica que utiliza luz láser. El láser se vislumbra como un posible sustituto para las lámparas actualmente utilizadas en los dispositivos de proyección de imágenes, principalmente por su saturación de color, su alta eficiencia en la generación de luz y su mayor vida útil.

El avance de los píxeles en pantalla viene de los tres láseres que emiten un rayo de cada uno de los tres colores primarios (rojo, verde y azul).

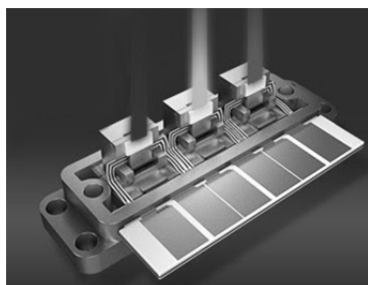


Figura 4-44. Iluminación de los píxeles: emisión un rayo de cada uno de los tres colores primarios.

El láser de color rojo se viene utilizando en aplicaciones informáticas y electrónicas desde hace tiempo, por lo que es una tecnología suficientemente probada e implementada. Sin embargo, no ocurre lo mismo en el caso de los láseres que emiten en las longitudes de onda azul y verde. La radiación de estos colores presenta una menor longitud de onda, y se sigue trabajando en el desarrollo de dispositivos semiconductores para conseguir buenos resultados en el campo de la imagen, ya que es importante que la potencia del rayo emitido sea la correcta.

La tecnología láser se basa en el chip **GLV** (*Grating Light Valve*) como pieza fundamental del sistema, por lo que también se la conoce como televisión GLV o GLV TV.

Se trata de una válvula conformada por seis rejillas recubiertas de un material reflectante, tres de ellas fijas y las otras restantes, móviles. Las tres rejillas móviles se mueven según el voltaje de entrada. Si a estas se les aplica el máximo voltaje, se alinean perfectamente con las tres rejillas fijas, difractando la máxima cantidad de la luz láser incidente (lo que correspondería al valor 255). En cambio, si el voltaje de entrada es el mínimo, se reflejará completamente ese color.

La luz láser de cada componente RGB difractada por la válvula GLV llega hasta una lente encargada de recogerla y de descartar la luz reflejada. Esta separación se consigue mediante un **filtro de Fourier**. Una vez que la luz difractada ha sido recolectada por la lente, se realiza la combinación final de los tres colores primarios RGB para formar el haz de luz definitivo que corresponderá al color del píxel. Antes de llegar al espejo de escaneado, la luz láser pasa por una **lente de proyección**, para darle la suficiente potencia y evitar problemas de visualización en pantalla que podrían ocurrir si la intensidad del haz fuera demasiado baja.

Necsel

SONY mostró un primer televisor láser en la feria del CES 2007, basado en tecnología láser TV de la compañía Novalux. Esta compañía creó la tecnología láser **Necsel**; un tipo de semiconductor láser que presenta las características ópticas y físicas necesarias para aplicaciones de visualización de imágenes.

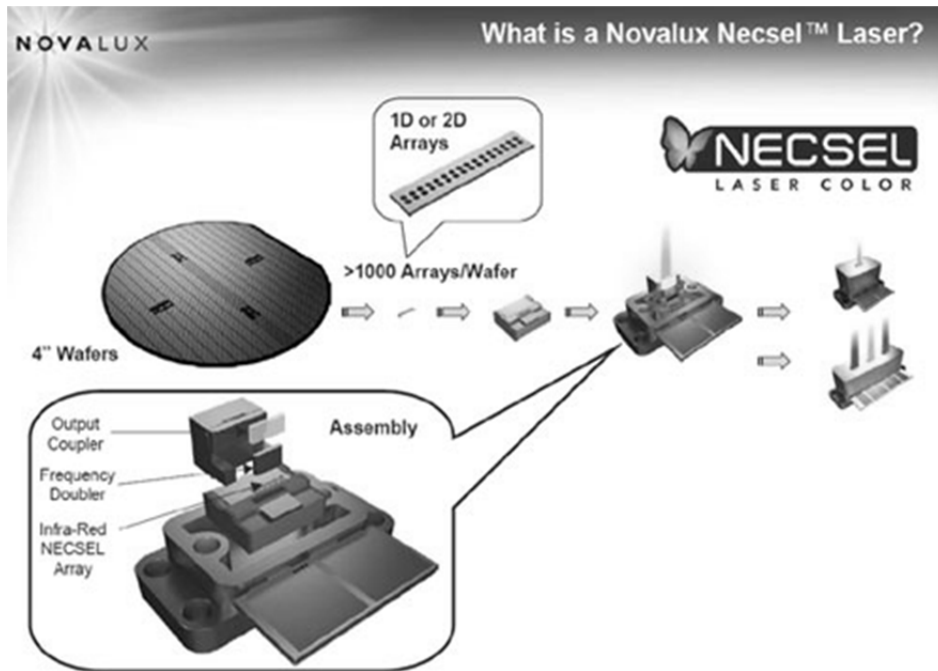


Figura 4-45. Estructura de un píxel con láser Necsel.

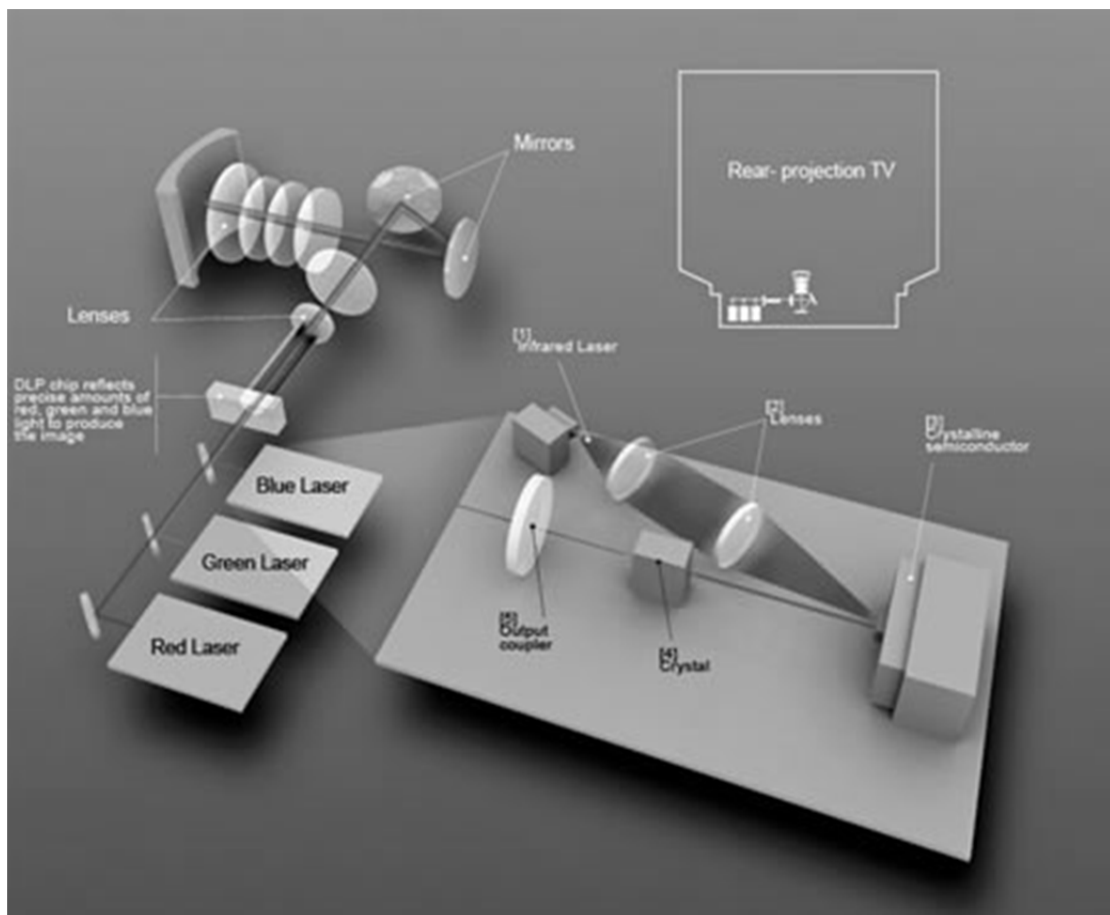


Figura 4-46. Esquema de proyección con la tecnología Necsel.

Según Novalux, sus pantallas de tecnología láser presentan las siguientes características:

-Proporcionan una paleta de colores más rica e intensa que las pantallas convencionales PDP, LCD y CRT.

- Tienen mitad de peso y coste que las pantallas PDP o LCD.
- Presentan un consumo aproximadamente un 75% menor que las pantallas PDP o LCD.
- Son tan delgadas como las pantallas PDP o LCD.
- Tienen una vida útil de más de 50.000 horas.



Figura 4-47. Comparativa Láser (izquierda) vs. Plasma (derecha).

4.3.9 Tecnología DLP

Digital Light Processing (en español "Procesamiento Digital de Luz") es una tecnología usada en proyectores y televisores de proyección. El DLP fue desarrollado originalmente por Texas Instruments, y sigue siendo el único fabricante de esta tecnología, aunque muchos productos de mercado autorizados están basados en sus Asics.

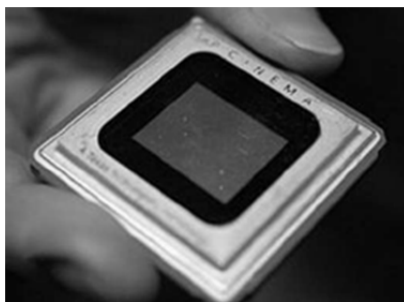


Figura 4-48. Tamaño de una matriz de visualización con tecnología DLP.

En los proyectores DLP, la imagen es creada por espejos microscópicos dispuestos en una matriz sobre un chip semiconductor, conocido como *Digital Micromirror Device* (DMD). Cada espejo representa un píxel en la imagen proyectada.

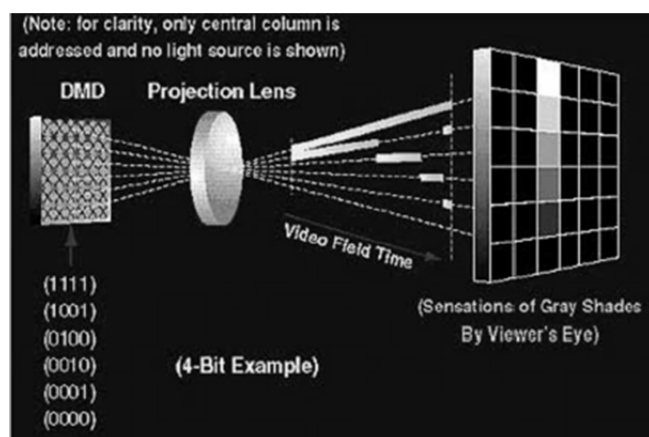


Figura 4-49. Ejemplo de proyección con 4 bits con un proyector DLP.

El número de espejos se corresponde con la resolución de la imagen proyectada: las matrices de 800×600, 1024×768, y 1280×720 son algunos de los tamaños comunes de DMD. Estos espejos pueden ser recolocados

rápidamente para reflejar la luz a través de la lente o sobre un disipador de calor. En la Figura 4-50 se muestra la estructura de los espejos reflectantes del DMD.

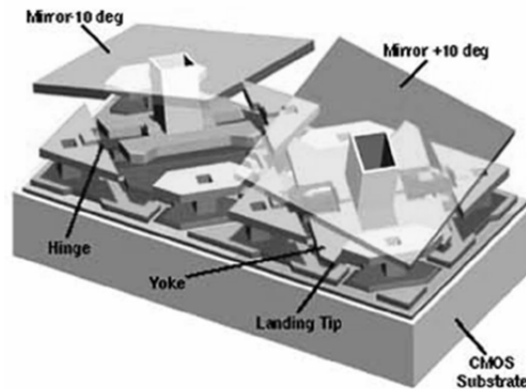


Figura 4-50. Estructura de los espejos reflectantes del DMD.

En un proyector con un sólo chip DMD, los colores son producidos colocando una rueda de color entre la lámpara y el DMD donde se refleja a través de la óptica. La rueda de color se divide generalmente en tres sectores, los colores primarios: rojo, verde y azul, y una sección clara adicional para el brillo, como se aprecia en la Figura 4-51. Puesto que el sector claro reduce la saturación del color, en algunos modelos puede ser inhabilitado con eficacia, y en otros se omite en conjunto.



Figura 4-51. Rueda de color.

En la Figura 4-52 se muestra el proceso de proyección con un solo chip. El chip DMD se sincroniza con el movimiento de rotación de la rueda de color para mostrar el componente verde en el DMD cuando la sección verde de la rueda de color está delante de la lámpara. Lo mismo ocurre con las secciones rojas y azules. Las imágenes rojas, verdes y azules se muestran así secuencialmente con una frecuencia suficientemente alta para que el observador vea la imagen compuesta en "color completo".

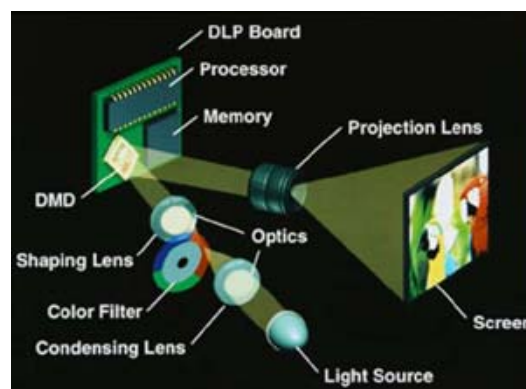


Figura 4-52. Proceso de proyección con un solo chip.

En los primeros modelos, había una rotación por marco. Los modelos más recientes hacen girar la rueda a una velocidad igual al doble de la frecuencia del marco, y algunos también repiten el patrón del color dos veces alrededor de la rueda, lo que significa que la secuencia se puede repetir hasta cuatro veces por marco.

Existe un efecto visual denominado **efecto arcoíris en DLP** se describe como flashes breves de "sombras" rojas/azules/verdes observadas principalmente cuando el contenido se caracteriza por objetos con brillo o blancos sobre un fondo oscuro o negro, como se muestra en la Figura 4-53 (los créditos del final de una película son un ejemplo común). Algunas personas perciben estos efectos de arco iris todo el tiempo, mientras que otras personas sólo los ven cuando mueven sus ojos a través de la imagen. Incluso otras personas no lo notan. El efecto se basa probablemente en el concepto del umbral de la fusión del parpadeo. En la Figura 51 se muestra cómo se ve una circunferencia blanca a través de una cámara mientras se mueve horizontalmente, con una exposición larga. La luz blanca está dividida en sus componentes de color. El efecto arcoíris se da cuando esto es visible a simple vista.

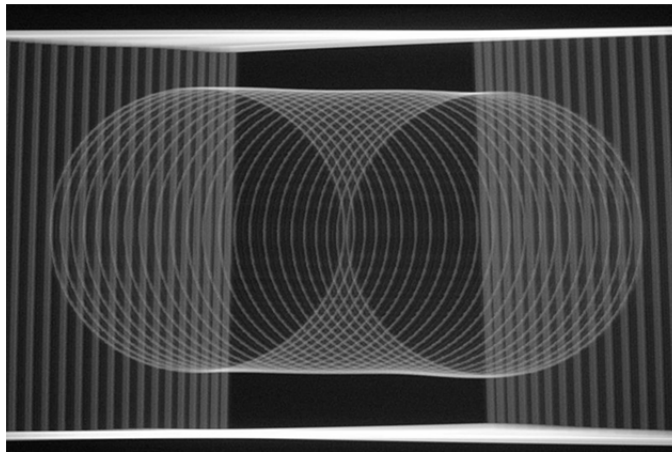


Figura 4-53. Efecto arcoíris.

Los fabricantes de sistemas de proyectores DLP de un chip utilizan las ruedas de color con una velocidad de rotación más alta, o con más segmentos de color para reducir al mínimo estos efectos.

Un proyector DLP de tres chips utiliza un prisma para dividir la luz de la lámpara, y cada uno de los colores primarios de la luz se encamina hacia su propio chip DMD, después se recombina y se dirige hacia el exterior a través de las lentes. Los sistemas DLP de un sólo chip son capaces de mostrar 16,7 millones de colores, mientras que los sistemas DLP de tres chips pueden mostrar hasta 35 billones de colores. Los proyectores de tres chips no sufren el efecto arcoíris, ya que los tres componentes del color (rojo, verde, y azul) se generan simultáneamente.

Entre las ventajas del DLP se pueden destacar imágenes suaves, sin saltos, buena profundidad y contraste del color y además no se marcan las imágenes en la pantalla. Como principal inconveniente se pueden señalar el efecto de arcoíris.

El principal competidor de DLP se conoce como **LCoS** (Liquid Crystal on Silicon), que crea las imágenes usando un espejo inmóvil montado sobre un chip, y utiliza una matriz de cristal líquido para controlar cuánta luz se refleja.

4.3.10 Tecnología LCoS

La tecnología LCoS podría considerarse como un híbrido entre las tecnologías, también de proyección/reproducción, LCD y DLP. De esta forma, para comprender el funcionamiento de LCoS es necesario entender primero el funcionamiento de LCD y DLP.

La tecnología LCoS utiliza una pantalla de cristal líquido y un dispositivo de silicio. La pantalla está aplicada sobre el dispositivo, de manera que éste aplica una corriente eléctrica sobre las partículas de cristal líquido y cambia su polarización para que éstas dejen o no pasar la luz de una lámpara. El dispositivo de silicio, además,

posee una superficie reflectante, de este modo la luz que atraviesa la pantalla incide sobre dicha superficie y es reflejada. Al aplicar diferentes niveles de voltaje sobre las partículas de cristal líquido se consiguen diferentes niveles de brillo intermedio entre luz y no-luz, y la cantidad de luz reflejada conforma la imagen. Así, podemos decir que la tecnología LCoS combina la técnica de transmisión, ya que las partículas de cristal líquido dejan (o no) pasar la luz, y la técnica de reflexión, ya que la superficie reflectante del dispositivo de silicio devuelve el haz de luz.

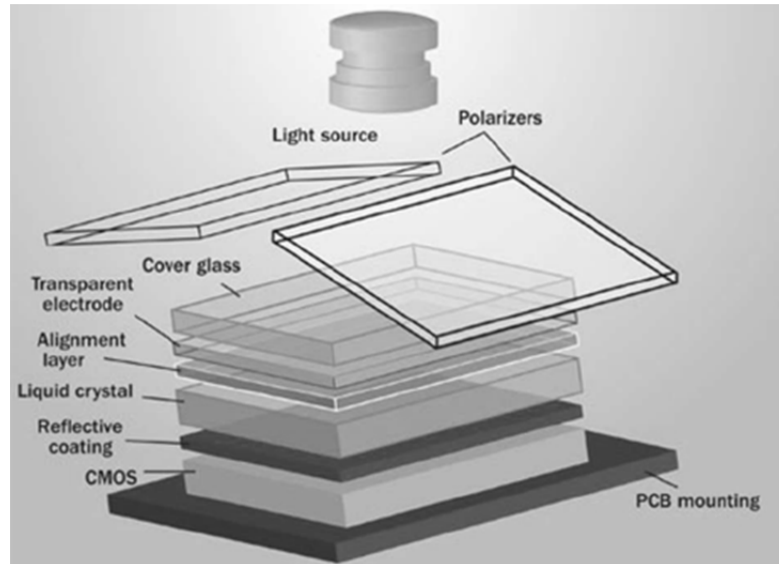


Figura 4-54. Estructura del LCoS.

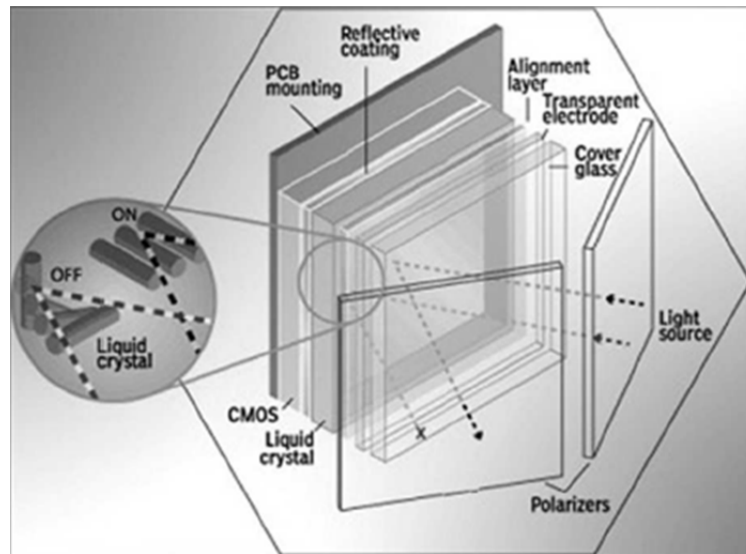


Figura 4-55. Detalle de la formación de la imagen en el LCoS.

Según la técnica que se utilice para representar los colores y su diseño podemos realizar una clasificación en dos grandes categorías: tecnologías con un único dispositivo LCoS y tecnologías con tres dispositivos LCoS. A continuación se explica su funcionamiento y características básicas de cada método, aunque cada fabricante posee sus propias particularidades.

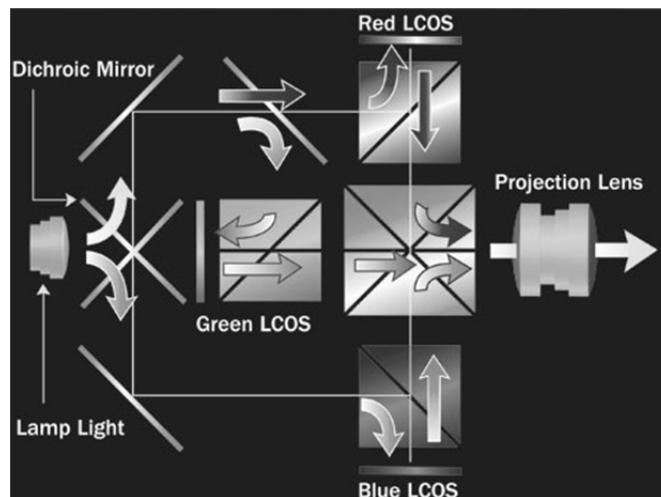


Figura 4-56. Proceso de proyección:

- 1°.- La lámpara crea un haz de luz blanca.
- 2°.- El haz pasa a través de una lente condensadora que concentra y dirige la luz.
- 3°.- La luz blanca se separa en roja, verde y azul.
- 4°.- Cada uno de los haces de color entra en contacto con uno de los tres microdispositivos LCoS - uno para el rojo, verde y azul.
- 5°.- Las luces reflejadas por los tres microdispositivos pasan a través de un prisma que las combina.
- 6°.- El Prisma dirige la luz hacia una lente de proyección para ampliar la imagen y mostrarla en la pantalla.

4.3.10.1 Tecnologías con tres dispositivos LCoS

Como su propio nombre indica, utilizan tres dispositivos LCoS (tres pantallas formadas por partículas de cristal líquido y sus correspondientes dispositivos de silicio) y cada uno de ellos se encarga de formar la imagen en uno de los tres colores primarios (rojo, verde y azul). Se basan en la combinación de diferentes espejos especiales (que dividen el haz de luz en dos, reflejando y dejando pasar la luz simultáneamente) para dividir y conducir la luz proveniente de la lámpara hacia los tres dispositivos distintos y para, posteriormente, integrar las tres imágenes formadas por cada uno de los tres dispositivos. De esta manera, mediante la combinación de las tres imágenes en los tres colores primarios se consigue formar la imagen en color real.

4.3.10.2 Tecnologías con un único dispositivo LCoS

Como su propio nombre indica, utilizan un único dispositivo LCoS (una única pantalla formada por partículas de cristal líquido y su correspondiente dispositivo de silicio). Se basan en la consecución de los distintos colores mediante filtros. En éste apartado podemos distinguir dos técnicas con diseños diferentes: la rueda de color, que es la más común, y la denominada “Color Dots” (matriz de color).

En el primer caso se utiliza un filtro en forma de rueda que se coloca frente a la luz proveniente de la lámpara, el cual (el filtro) se encuentra dividido en tres secciones, una para cada componente del color (los tres colores primarios: rojo, verde y azul). El dispositivo de silicio, que es el responsable de aplicar las diferentes tensiones eléctricas a las partículas de cristal líquido para cambiar su polarización y que éstas dejen pasar más o menos cantidad de luz, se sincroniza con la rueda de color, de manera que cuando la sección roja de la rueda de color se encuentre frente al haz de luz proveniente de la lámpara, el dispositivo LCoS emita la cantidad de luz correspondiente al color rojo, y así sucede también con las otras dos secciones correspondientes a los otros dos

colores primarios restantes de la rueda de color. De este modo, se muestran tres imágenes, una roja, una verde y una azul (una para cada color primario, las componentes necesarias para crear el resto de colores) a una frecuencia elevada tal que el observador visualiza una imagen en color real formada a partir del resultado de que éste integre las tres imágenes en una sola. Al solo disponer de un solo LCoS, también sufre el efecto arcoíris.

Ventajas:

- Consiguen resoluciones muy altas con el mínimo de elementos (por ejemplo, en el caso de la tecnología de proyección/reproducción DLP, para conseguir mayores resoluciones son necesarios chips DMD formados por un mayor número de micro espejos), debido a la poca distancia existente entre píxeles, de modo que se consigue una imagen mucho más natural y el pixelado es prácticamente inexistente.
- Los bordes de sus píxeles se muestran más suavizados, en contraste con los bordes afilados de los microespejos del sistema DLP, de manera que reproduce una imagen de un modo mucho más suave y no con un aire tan sintético y artificial, que puede provocar una sensación de imagen más forzada, menos natural.
- Consigue mejores colores saturados que LCD y DLP.

Inconvenientes

- Tecnología más cara que otras como LCD y DLP.
- Los proyectores/reproductores que utilizan esta tecnología son más pesados y menos compactos (sobre todo en el caso de utilizar tecnologías con tres dispositivos LCoS).
- Su nivel de negro es pobre, por lo que no consigue un contraste tan elevado como otras tecnologías.
- La vida de las lámparas utilizadas para este sistema es menor y además poseen un coste elevado.
- Compite con los proyectores/reproductores DLP para ser la tecnología utilizada en la llegada del cine digital, pero de momento se le considera inferior a ésta, ya que una de las aplicaciones que sitúan al cine digital sobre el cine analógico es la reproducción en 3D, y debido a las altas frecuencias de cuadro necesarias para su reproducción la tecnología LCoS no está capacitada para ello.
- Elevado coste económico.

Aplicaciones

Uso doméstico como proyector/reproductor, en la proyección/reproducción de cine digital, interruptores y conmutadores ópticos, almacenamiento de datos holográficos y proyección/reproducción de material holográfico.

4.4 Cuestionario

1-La resolución de un monitor:

- a) Es la cantidad de puntos en sentido horizontal y vertical que se tiene de una imagen.
- b) Es la cantidad de puntos en sentido horizontal y vertical que se pueden presentar en pantalla (número de puntos por unidad de longitud).
- c) Determina el nivel de intensidad de iluminación de fondo en la imagen presentada en pantalla.
- d) Ninguna de las anteriores es cierta.

2-Asociar cada señal con el conector que ésta usa.

- | | |
|---------------------|--|
| a) Vídeo compuesto. | 1-MinDIN-4. |
| b) S-Vídeo. | 2-Conector específico de este formato de 24 pines. |
| c) Componentes. | 3-Conector específico de este formato de 19 pines. |
| d) VGA. | 4-RCA-Amarillo. |
| e) HDMI. | 5-SUB-D15. |
| d) DVI. | 6-RCA para cada señal. |

3- Existe cierta similitud entre los monitores CRT y los SED, mientras los primeros sólo usan un cañón de electrones para representar las imágenes, los monitores SED usan un emisor de electrones para cada fila de píxels.

- a) Verdadero.
- b) Falso.

4- Los monitores con tecnología LCD se basan:

- a) En la emisión de luz de muy poca intensidad, por lo que el suministro eléctrico que requieren es mínimo.
- b) Uso de cristales líquidos para el bloqueo de la luz.
- c) Un haz de rayos catódicos constante dirigido contra una pantalla de vidrio recubierta de fósforo y plomo.
- d) Un dispositivo de pantalla plana habitualmente usada en televisores que consta de una matriz de celdas, cada una de las cuales se componen a su vez de tres sub-celdas «sub-píxeles», correspondientes a los colores rojo, verde y azul, situadas entre dos paneles de cristal que contienen una mezcla de gases nobles (neón y xenón).

5- Los cristales líquidos de tipo esmétrico siempre tienen los ejes moleculares orientados a una misma dirección y sus centros de masa ordenados unos respecto a otros.

- a) Verdadero.
- b) Falso.

6-El brillo en los monitores:

- a) Es la diferencia relativa de intensidad luminosa entre los puntos de una imagen.
- b) Es luz policromática que contiene diversos haces monocromáticos con distintas longitudes de onda
- c) Determina el nivel de intensidad de iluminación de fondo en la imagen presentada en pantalla
- d) Ninguna de las anteriores es cierta.

7-El contraste dinámico es una técnica, usada en dispositivos LEDs o Plasma, que modifica intencionalmente el brillo en determinadas imágenes para tratar de optimizar artificialmente el contraste.

- a) Verdadero.
- b) Falso.

8-El monitor de Plasma o PDP:

- a) Es un dispositivo de pantalla LCD habitualmente usado en televisores que consta de una matriz de celdas, cada una de las cuales se componen a su vez de tres sub-celdas «sub-píxeles», correspondientes a los colores rojo, verde y azul, situadas entre dos paneles de cristal que contienen una mezcla de gases nobles (neón y xenón).
- b) Es un dispositivo de pantalla plana habitualmente usado en televisores que consta de una matriz de celdas, cada una de las cuales se componen a su vez de tres sub-celdas «sub-píxeles», correspondientes a los colores rojo, verde y azul, situadas entre dos paneles de cristal que contienen una mezcla de gases nobles (neón y xenón).
- c) Es un dispositivo compuesto de paneles o módulos de LED (diodos emisores de luz) debidamente compuestos por LEDs RGB. Se utilizan los LEDs disponiéndolos en forma de matriz utilizando diodos de distintos colores RGB para formar el píxel, de tal forma que se pueden obtener pantallas de LED de diversos tipos como lo serían: tipo indicadores, informativas, publicitarias y de alta resolución de vídeo a todo color
- d) Ninguna de las opciones anteriores es cierta.

9-Los paneles SED se caracterizan por usar la misma tecnología CRT para cada uno de los campos de la imagen.

- a) Verdadero.
- b) Falso.

10- La tecnología láser:

- a) Tecnología usada en proyectores y televisores de proyección.
- b) Es una tecnología de visualización y proyección de vídeo basada en optoelectrónica que utiliza luz láser. El láser se vislumbra como un posible sustituto para las lámparas actualmente utilizadas en los dispositivos de proyección de imágenes, principalmente por su saturación de color, su alta eficiencia en la generación de luz y su mayor vida útil.
- c) Está altamente difundida en la actualidad. Es muy útil porque permiten mostrar información o datos de manera muy clara. La mayoría de los electrodomésticos y diversos equipos electrónicos traen uno o varios de ellos porque presentan la gran ventaja de consumir poco.
- d) Ninguno de las opciones anteriores es cierta.

11- Los monitores de matriz activa:

- a) Al presentar elementos de imagen excitados en forma directa, sin interferir en los demás, hace que el conexionado sea muy complejo, llegando a tener en pantallas millones de conexiones, sin posibilidad de reducir su número.
- b) Están formados por dos filtros polarizadores con filas de electrodos transparentes situados perpendicularmente por encima y por debajo del LCD.
- c) Al presentar elementos de imagen excitados en forma directa, sin interferir en los demás, hace que el conexionado sea muy complejo, pero se puede reducir direccionando los píxeles en filas y columnas, lo que reduce el número de conexiones de millones a miles.

d) Ninguna de las anteriores es cierta.

12-Los monitores de matriz pasiva:

a) Al presentar elementos de imagen excitados en forma directa, sin interferir en los demás, hace que el conexionado sea muy complejo, llegando a tener en pantallas millones de conexiones, sin posibilidad de reducir su número.

b) Están formados por dos filtros polarizadores con filas de electrodos transparentes situados perpendicularmente por encima y por debajo del LCD.

c) Al presentar elementos de imagen excitados en forma directa, sin interferir en los demás, hace que el conexionado sea muy complejo, pero se puede reducir direccionando los píxeles en filas y columnas, lo que reduce el número de conexiones de millones a miles.

d) Ninguna de las anteriores es cierta.

13-La base física del funcionamiento de los LCD es el cambio de la dirección del eje óptico bajo la acción de un campo eléctrico.

a) Verdadero.

b) Falso.

14- Un diodo orgánico de emisión de luz:

a) Es un tipo de dispositivo visualizador para pantallas planas caracterizado por usar la tecnología de las pantallas de tubo tradicionales (CRT) para cada uno de los puntos (píxeles) mostrados en pantalla pero con el tamaño de los monitores LCD

b) Se basa en el mismo funcionamiento que las pantallas de plasma: cada píxel emite su propia luz

c) Es una tecnología altamente difundida en la actualidad. Es muy útil porque permiten mostrar información o datos de manera muy clara. La mayoría de los electrodomésticos y diversos equipos electrónicos traen uno o varios de ellos porque presentan la gran ventaja de consumir poco.

d) Es un diodo que se basa en una capa electroluminiscente formada por una película de componentes orgánicos que reaccionan, a una determinada estimulación eléctrica, generando y emitiendo luz por sí mismos.

15-La resolución de una imagen es la cantidad de puntos en sentido horizontal y vertical que se pueden presentar en pantalla (número de puntos por unidad de longitud).

a) Verdadero.

b) Falso.

16- La resolución de la pantalla es la cantidad de puntos en sentido horizontal y vertical que se pueden presentar en pantalla (número de puntos por unidad de longitud).

a) Verdadero.

b) Falso.

17-De los 19 pines que consta el conector HDMI, únicamente 2 (los pines 14 y 15) son los que llevan la información de vídeo.

a) Verdadero.

b) Falso.

18-Indica el tipo de cristal líquido usado para hacer pantallas:

a) Colestéricos

- b) Nemático.
- c) Dicroicos.
- d) No se usa cristal líquido para hacer pantallas.

19- Indica los tipos de direccionamiento de iluminación de píxeles:

- a) Mediante matrices pasivas y activas.
- b) Mediante matrices activas y redundantes
- c) Mediante memorias.
- d) No es necesario el direccionamiento, la captación de una cámara lleva consigo la información necesaria sobre qué pixel hay que iluminar.

20- Un proyector de vídeo o vídeo proyector es un aparato que recibe una señal de vídeo y proyecta la imagen correspondiente en una pantalla de proyección usando un sistema de lentes, permitiendo así mostrar imágenes fijas o en movimiento.

- a) Verdadero.
- b) Falso.

TEMA 2

Señales y formatos de audio y vídeo

5 DIGITALIZACIÓN DE LAS SEÑALES DE AUDIO Y VÍDEO

Las fuentes de audio y vídeo en origen son analógicas. Estas señales analógicas se pueden digitalizar para conseguir un mejor procesado, transmisión, almacenamiento, etc. En su destino, en la mayoría de los casos, volverán a convertirse a señales analógicas para su representación; por lo que la conversión A/D/A (analógico/digital/analógico) deberá introducir el menor error posible.

Las etapas de la digitalización son las que se muestran en la :

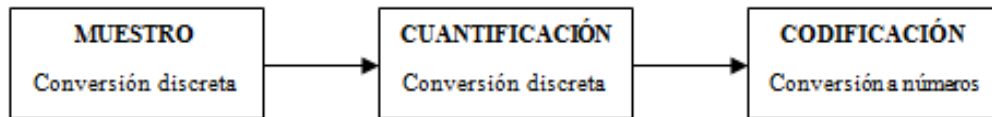


Figura 5-1. Etapas de la digitalización.

5.1 Muestreo

Cualquier digitalización está determinada por el Teorema de Muestreo o Nyquist, que define la relación entre el ancho de banda la señal a muestrear y la frecuencia de muestreo. Para obtener una señal discreta en tiempo que sea la versión fiel de una señal analógica, la frecuencia de muestreo f_s se obtiene con la expresión (5-1):

$$f_s = 2 \cdot BW \text{ Hz. (5-1)}$$

siendo BW el ancho de banda o frecuencia máxima de la señal a digitalizar.

En caso contrario, la señal discreta en tiempo incluye componentes espectrales no existentes en la señal analógica, no pudiéndose recuperar ésta sin error. Este error se denomina **aliasing**.

Para evitar este error se coloca el filtro antialiasing. Es un filtro paso banda que limita el ancho de banda de la señal a digitalizar. Como cualquier filtro paso bajo introduce distorsiones lineales en la señal:

- Distorsión de amplitud al no ser plana la banda de paso.
- Distorsión de fase o retardo de grupo.

El filtro antialiasing no es ideal por lo que puede aparecer cierto aliasing. Hay que llegar a un compromiso entre el aliasing y las características del filtro que determinan la distorsión (pendiente, frecuencia de corte, rizado,...).

5.2 Cuantificación y codificación

El margen dinámico de la señal se cuantifica en 2^n valores posibles, siendo n el número de bits que se utilizan para codificar cada muestra. Luego aparece el **error de cuantificación** que se apreciará al recuperar la señal analógica. El error de cuantificación es la diferencia entre el valor real de la muestra y el valor cuantificado.

Los valores de señal cuantificados tomarán los siguientes valores binarios codificados:

- De 0 a $2^n - 1$, para señales con valores positivos.
- Para señales con valores positivos y negativos se puede utilizar dos métodos para representar los valores negativos:
 - De $-(2^{(n-1)} - 1)$ a $2^{(n-1)} - 1$, método de **bit de signo**: el bit más significativo MSB determina el signo del valor expresado ("0" para el + y "1" para el -) y los bits siguientes la magnitud. Hay dos ceros.
 - De $-2^{(n-1)}$ a $2^{(n-1)} - 1$, método de **complemento a 2**: igual que antes pero para los valores negativos la magnitud se representa en complemento a dos. El complemento a dos se calcula invirtiendo los bits y sumando 1.

Por ejemplo si $n = 8$ bits:

- De 0 a 255
- De -127 a 127 con bit de signo, -127 = 1111 1111
- De -128 a 127 con complemento a dos, -128 = 1000 0000 y -1 = 1111 1111

5.3 Digitalización de audio

El audio digital es una tecnología muy sofisticada y aunque sus ideas principales fueron concebidas en los años 20, no fue hasta 1970 cuando se empezó a comercializar con ella. Es decir, se necesitaron 50 años de desarrollo tecnológico para lanzar definitivamente el audio digital. Las técnicas digitales usadas en grabación, reproducción, almacenamiento, procesado y transmisión de señales digitales de audio involucran conceptos que escapan a las técnicas analógicas del sonido.

5.3.1 Muestreo de la señal de audio

El ancho de banda básico de la señal de audio es 20 kHz. En la Figura 5-2 se muestra el esquema básico de la digitalización de audio.

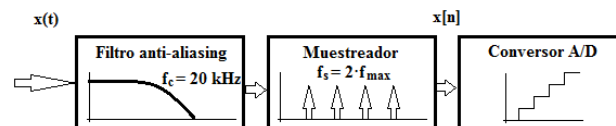


Figura 5-2. Esquema básico de un sistema de digitalización.

En la Figura 5-3 se resume el proceso de muestreo y recuperación de una señal analógica.

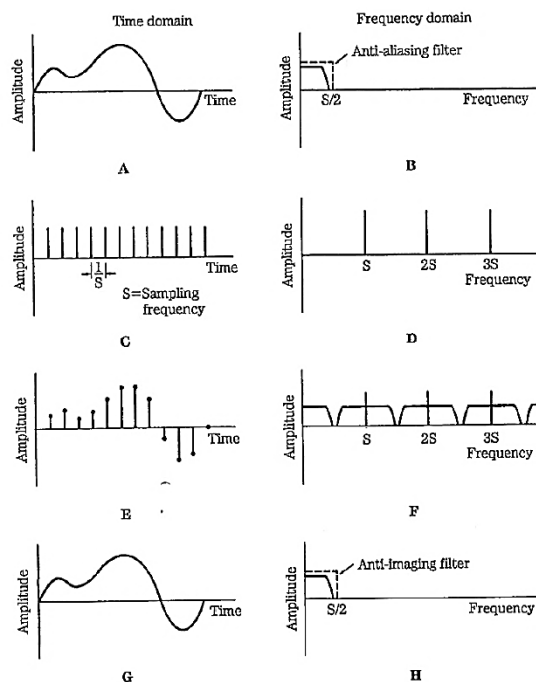


Figura 5-3. Dominio del tiempo (izqda.) y dominio de la frecuencia (dcha.). A: La señal de entrada después del filtro anti-aliasing. B: Espectro de la señal de entrada. C: Señal de muestreo. D: Espectro de la señal de muestreo. E: Señal de entrada muestreada. F: Espectro de la señal de entrada muestreada. G: Salida de la señal después del filtro paso bajo recuperador. H: Espectro de la señal de salida.

Los fabricantes han determinado una frecuencia de muestreo de 44,1 kHz para el CD por su tamaño, tiempo de reproducción y coste medio. Para equipos de estudio y grabadores se utiliza 48 kHz. Estas frecuencias hacen que el filtro antialiasing tenga mucha pendiente y por lo tanto introduzca mucha distorsión en la señal digitalizada.

Por otro lado, se utilizan frecuencias de 96 kHz. En DVD-Audio y en discos Blu-ray se emplean frecuencias de muestreo de hasta 192 kHz. Los filtros antialiasing serán menos abruptos y por lo tanto con menor distorsión.

El muestreo real no se hace con deltas sino con pulsos debido al funcionamiento de los circuitos de muestreo y retención (*sample&hold*). La amplitud de la señal analógica se va integrando durante un tiempo τ , proporcionando un valor cada τ segundos; que equivale a muestrear con un pulso de ancho τ . El efecto es que el espectro de la señal muestreada queda multiplicado por una sinc cuyo primer nulo está en $1/\tau$ resultando que existe un cierto filtrado paso bajo. Luego por lo tanto, cuanto menor τ menor el efecto, o sea cuanto mayor sea la frecuencia de muestreo. Este efecto del circuito de muestreo y retención se llama **efecto apertura**.

5.3.2 Cuantificación de la señal de audio

El muestreo y la cuantificación son los dos pilares fundamentales de la digitalización de audio, respectivamente determinan el ancho de banda y la resolución del sistema. Debido a la cuantificación aparece una limitación en la resolución de la digitalización, lo que afecta a la precisión con la que se toman los valores de cada pulso a la salida del muestreador. La precisión de cuantificación depende principalmente de la longitud de la palabra digital que se vaya a utilizar, es decir, del número de bits que representa cada muestra. Al representar con un número finito de escalones los infinitos valores de una forma de onda analógica se produce el **error de cuantificación**, similar al ruido característico de los sistemas analógicos. La diferencia con el ruido analógico es que el error de cuantificación varía según la amplitud de la forma de onda.

Existen dos tipos de cuantificaciones.

a) Cuantificación uniforme o lineal

La amplitud de una señal analógica es cuantificada con un número N finito de escalones de cuantificación, todos ellos de igual tamaño.

Si se tiene una palabra de n bits se podrán obtener $N = 2^n$ niveles de cuantificación. En la Tabla 5-1 se ejemplifican algunos de estos valores.

Tabla 5-1. Niveles posibles según el número de bits usados.

$2^1 = 2$	$2^{13} = 8192$
$2^2 = 4$	$2^{14} = 16384$
$2^3 = 8$	$2^{15} = 32768$
$2^4 = 16$	$2^{16} = 65536$
$2^5 = 32$	$2^{17} = 131072$
$2^6 = 64$	$2^{18} = 262144$
$2^7 = 128$	$2^{19} = 524288$
$2^8 = 256$	$2^{20} = 1048576$
$2^9 = 512$	$2^{21} = 2097152$
$2^{10} = 1024$	$2^{22} = 4194304$
$2^{11} = 2048$	$2^{23} = 8388608$
$2^{12} = 4096$	$2^{24} = 16777216$

Como se aprecia en la Tabla 5-1, cada vez que se añade un bit a la palabra de audio el número de niveles que es posible representar se multiplica por dos. Luego a mayor número de bits, mayor aproximación, es decir el error de cuantificación disminuye pero no llega a desaparecer.

Hay un momento en el que el error de cuantificación se vuelve imperceptible. La mayoría de los fabricantes establecen un valor de 16 a 20 bits para hacer una representación adecuada. Por ejemplo en formatos DVD y Blu-ray se usan palabras de 24 bits y técnicas avanzadas de eliminación de ruido de cuantificación como el *dithering* y *noise shaping*. La técnica del *dithering* se explica en el ANEXO A. Se puede demostrar que aproximadamente la relación entre la señal y el ruido de cuantificación varía según la ecuación dB. (5-2):

$$S/E = 6,02 \cdot n + 1,76 \text{ dB. (5-2)}$$

Luego cada vez que se añade un bit a la palabra digital de audio, la relación S/E aumenta 6 dB, o lo que es lo mismo, se multiplica por dos. En el ANEXO A se detalla la obtención de la ecuación dB. (5-2).

El error de cuantificación se traduce en distorsión de la señal de audio. Si la señal tiene nivel alto la percepción de esta distorsión será prácticamente inaudible, en cambio si se trabajan con niveles bajos de señal, el error puede llegar a percibirse.

b) Cuantificación no uniforme

Se usan diferentes amplitudes de escalones de cuantificación. Se eligen pequeños escalones para tensiones de pequeñas y escalones grandes para tensiones grandes. Con esto se consigue mantener el número total de niveles y el ancho de banda de transmisión pero se reduce el error de cuantificación, haciéndolo dependiente del nivel de la señal de entrada.

Para realizar esa distribución no uniforme de escalones-niveles se recurre a la compresión y expansión de la señal tratada tal y como se muestra en la Figura 5-4.

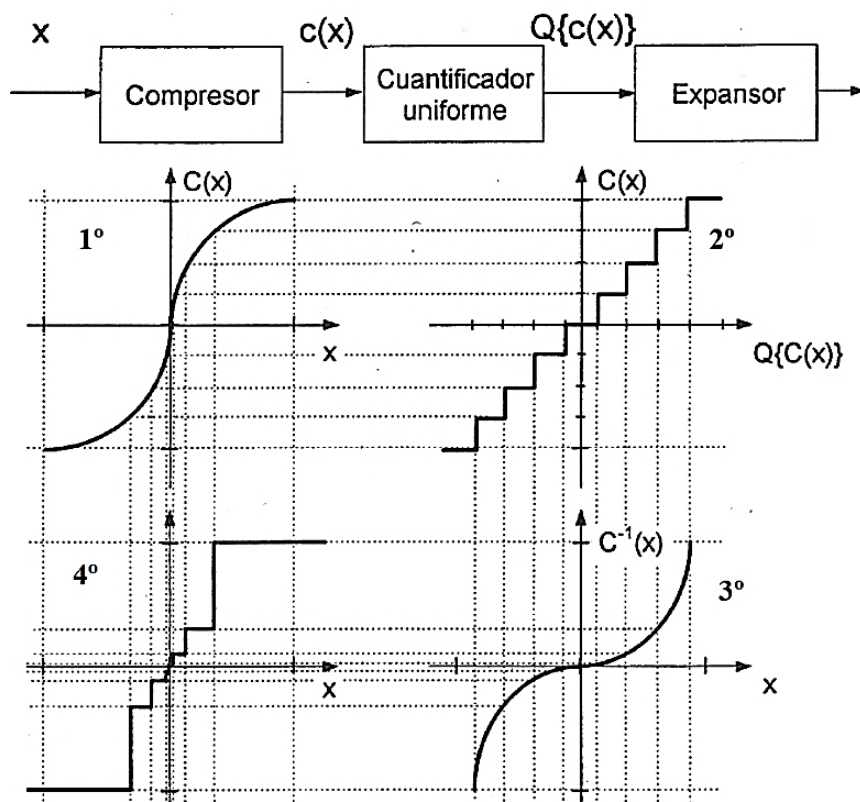


Figura 5-4. Proceso de cuantificación no uniforme. La señal de entrada se comprime para que al pasar por un cuantificador uniforme el resultado sea como si se hubiese contado con un cuantificador no uniforme en el proceso.

Con este procedimiento se puede obtener la expresión (5-3):

$$S/E_{\text{no uniforme}} = G^2 S/E_{\text{uniforme}} \quad (5-3)$$

Donde la ganancia de compresión, G , vendrá dada por el tipo de ley de compresión utilizado. Existen dos leyes de compresión típicas, usada ampliamente en la digitalización de voz: la ley A y la ley μ .

a) Ley A

Tiene una característica lineal en las proximidades del origen ($|x| \leq x_{\text{max}}/A$) y otra logarítmica en el resto. El valor que da a la ganancia respecto a la cuantificación uniforme se obtiene con la expresión (5-4):

$$G = \frac{A}{1 + \ln(A)} \quad (5-4)$$

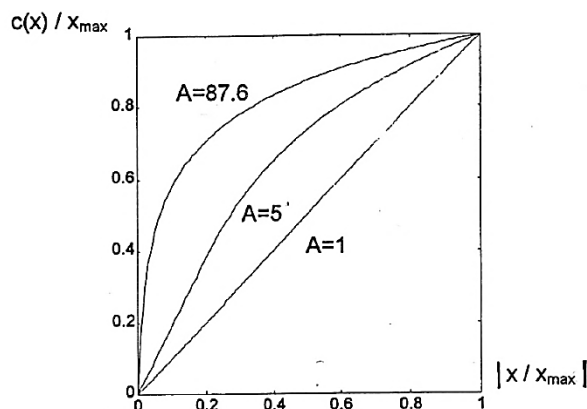


Figura 5-5. Representación de la Ley A con diferentes valores para su parámetro A.

b) Ley μ

Su ganancia viene dada por la expresión (5-5):

$$G = \frac{\mu}{\ln(1+\mu)} \quad (5-5)$$

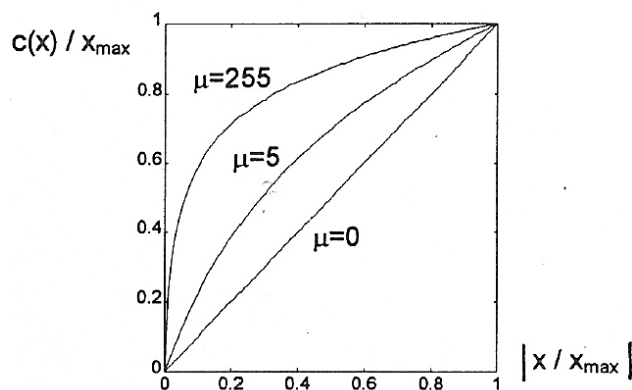


Figura 5-6. Representación de la Ley μ con diferentes valores para su parámetro μ .

5.3.3 Señal digital de audio AES3

La señal utilizada para transmitir señales digitales de audio se estandarizó en la interfaz AES3 también conocida en Europa como AES/EBU (*Audio Engineering Society/European Broadcasting Union*). Este interfaz es el usado en los equipos de estudio y de él se deriva el interfaz doméstico SPDIF.

La señal transmitida tiene las siguientes características:

- No tiene compresión.
- 2 canales de audio: típicamente estéreo L y R, con multiplexación TDM.
- Frecuencia de muestreo: 44.1 o 48 kHz.
- Número de bits por muestra: entre 16 y 24 bits.

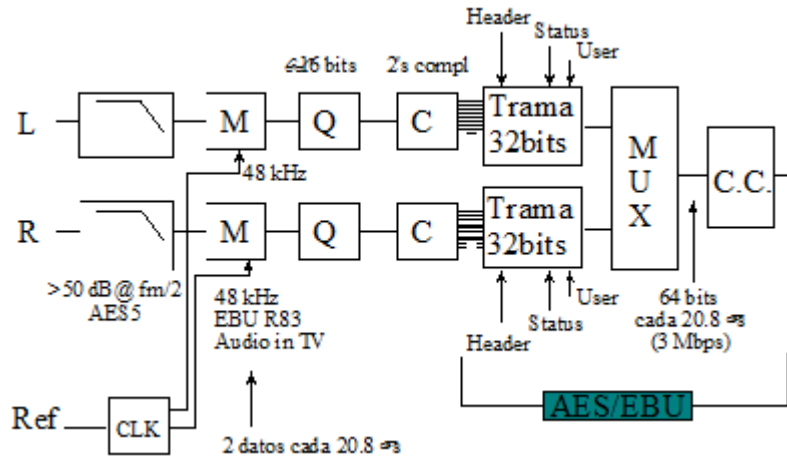


Figura 5-7. Diagrama de bloques de la digitalización de audio AES/EBU.

Los datos de cada muestra de cada canal digitalizado (hasta 24 bits) se introducen en una trama de datos digitales de 32 bits. El resto de bits (8) se usan para sincronizar y llevar información adicional al equipo que recibe la señal, p.e. la frecuencia de muestreo, el n° de bits por muestra. El régimen binario a la salida se obtiene de la expresión (5-6):

$$R_b = f_s \cdot 2 \cdot n = 48000 \cdot 2 \cdot 32 = 3,072 \text{ Mb/s. (5-6)}$$

El **código de canal** es la manera de representar físicamente (en tensiones) los símbolos que se transmiten. Cada símbolo transmitido corresponde a un bit ("1" o "0") o a una combinación de un conjunto de n bits. Estos símbolos o formas de onda de tensión son los que se transmiten por el cable. En la Figura 5-8 se puede ver la forma de onda de onda AES3.

El más simple y típico es el NRZ-L (*Non Return to Zero – Level*); donde el "1" se asocia a un nivel tensión mantenido toda la duración del símbolo. El "0" se asocia a otro nivel tensión distinto.

En el caso del interfaz AES/EBU se usa el código de canal BiPhase-Mark que es más adecuado para transmisión serie por cable porque con él se puede recuperar el reloj de sincronización mejor.

El código BiPhase-Mark es como sigue:

- Para cada nuevo bit, el nivel de tensión cambia al principio del símbolo.
- "0" el nivel se mantiene toda la duración del bit.
- "1" el nivel cambia a la mitad de la duración del bit.

Por lo tanto el nivel de tensión no determina el dato, sino la forma de onda del símbolo. El resultado es una señal cuyo pulso más estrecho dura la mitad del tiempo de bit, luego el $T_{\text{Símbolo}}$ es la mitad que el periodo de bit y el Régimen de símbolo o velocidad de transmisión R_s es el doble que el régimen binario, así como el reloj del canal es el doble que el reloj de bit.

El R_s 6,144 MBaudios, aunque sólo transporta 3,072 Mb/s.

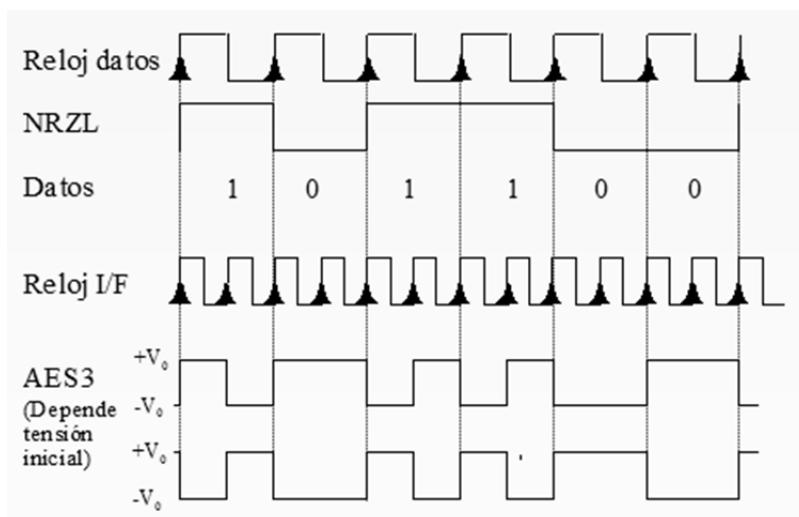


Figura 5-8. Forma de onda de la señal AES/EBU.

La estructura de este interfaz se ha ampliado para transportar otras variantes de la señal de audio digital como son: 1 canal a 96 kHz, señal comprimida MPEG o Dolby AC-3, canales de sistemas multicanal,... Para ello se estructuran tramas de 24 bytes (192 bits) con datos de audio y metadatos de información, corrección de errores, etc.

El **SP-DIF** (Sony/Philips *Digital Interface Format*) S/PDIF es una versión del protocolo estándar AES/EBU adaptada para aplicaciones comerciales, presentando pequeñas diferencias que lo hacen menos costoso a la hora de producir los componentes finales. Viene definido por la IEC 60958. La aplicación de consumo (SPDIF) se indica en el “channel status” con el primer bit del primer byte.

La diferencia de significado y análisis de muchos de los datos de la trama hacen que sea incompatible la aplicación de consumo con la profesional, por lo que AES3 y SPDIF son incompatibles aunque su contenido de audio sean iguales.

5.4 Digitalización de vídeo

En los centros de producción ya se viene trabajando desde hace bastante tiempo con señales digitales, aunque posteriormente esas señales hubieran de convertirse a señales analógicas para su posterior distribución. Actualmente ya no es necesaria esa conversión ya que se trabaja directamente con señales digitales en la transmisión.

Es cierto que en el proceso de generación de señales (en las cámaras) se siguen usando técnicas analógicas, por lo que será necesaria su inmediata digitalización para continuar con el proceso digital. En el futuro es previsible que se puedan obtener directamente las tres señales correspondientes a cada píxel independientemente de las tres señales correspondientes a los píxeles vecinos, por lo que el proceso de muestreo que hay que implementar en la actualidad, no será necesario llevarlo a cabo.

Dado que las señales digitales han de ser entendidas y procesadas por otros equipos, ha sido necesario normalizar la forma de obtenerlas, con objeto de que no haya ambigüedad en el proceso y se conozcan perfectamente sus características.

Por tanto, el proceso de obtención de señales digitales a partir de señales analógicas ha sido recogido en la Recomendación UIT-R BT 601 y recibe el título de *Parámetros de Codificación de Televisión Digital para estudios con formatos de imagen normal 4:3 y de pantalla ancha 16:9*, este proceso es el que se conoce popularmente como Norma Digital de Estudio, puesto que es la utilizada en todos los estudios de televisión para la obtención de sus señales digitales.

Esta norma es la que se desarrollará en este capítulo, puesto que este tipo de señales serán las utilizadas para la posterior comprensión y adecuación a la distribución por los diferentes medios de transmisión existentes, tanto por satélite como por cable o incluso en la distribución terrenal.

Los motivos que han llevado al establecimiento de esta Norma Digital de Estudio se pueden resumir en los siguientes puntos:

- a) Intentar que existan el mayor número de parámetros comunes a los Sistemas que trabajan con 525 líneas (sistema adoptado en algunos países como EE.UU) o con 625 líneas (sistema adoptado en España), puesto que esto representa una clara ventaja para los fabricantes de equipos, abaratando su precio; así como simplificará el necesario intercambio de programas entre diferentes organismos.
- b) Lograr que la Norma no sea rígida, sino que permita posibles ampliaciones futuras, con vista a los previsibles incrementos de calidad que serán demandados en el futuro.
- c) Trabajar con señales en componentes, decantadas como las más adecuadas para el tratamiento de señales de vídeo, al aprovechar las características del ojo humano. Como se verá posteriormente, las señales componentes podrán ser tanto luminancia y crominancia, como directamente las señales RGB, dependiendo de la necesidad.

5.4.1 Señales digitalizadas

Como se acaba de indicar, las señales analógicas disponibles pueden acceder al digitalizador en diferentes formatos. Básicamente son dos las posibilidades que se pueden plantear:

a) Trabajar con señales en componentes RGB

En principio, sin ninguna consideración adicional, ésta es la opción más correcta, puesto que sería digitalizar directamente las tres señales generadas en cada uno de los píxeles, sin ninguna manipulación inicial. Sin embargo, actuando de esta forma se sufren una serie de inconvenientes, que pueden concretarse principalmente en dos:

- Sería necesario digitalizar cada una de las tres señales con elevadas frecuencias de muestreo, puesto que no sería posible ningún tipo de ahorro de bits en ninguna de ellas. Cualquier defecto existente en alguna de las tres, se manifestaría como una degradación de la imagen, por lo que sería necesario un tratamiento muy preciso para todas y cada una de las tres señales, con el consiguiente régimen binario obtenido muy elevado.
- Habría que ser muy cuidadoso con los retardos existentes entre cada uno de los tres canales, puesto que cualquier pequeño desajuste en este aspecto sería desastroso para la visualización de la imagen.

Normalmente no se usan las señales RGB en la digitalización ya que su tratamiento es demasiado complejo.

b) Trabajar con señales de luminancia y diferencia de color

Esta es la opción que elegida, porque presenta innumerables ventajas, siendo la principal el hecho de que el ojo humano es muy sensible a la señal de luminancia y menos exigente con las señales de crominancia. Esto hace pensar que la señal de luminancia deberá ser tratada con mucha definición (frecuencias de muestreo altas), pero las señales de crominancia podrán ser tratadas con menos detalle, ya que el ojo humano no demanda para estas señales unas frecuencias de muestreo tan elevadas. Con esto se consigue un gran ahorro de bits, sin que la calidad de la imagen resulte perjudicada.

Se ha de especificar que las señales que se van a tratar están gamma-correctas y por lo tanto en su nomenclatura (según la Norma 601) se debería hacer con tilde, pero para simplificar las expresiones aparecerán sin ella.

La señal de luminancia que se va a utilizar obedece a la expresión (5-7):

$$Y = 0,299R + 0,587G + 0,114B. (5-7)$$

las señales diferencia de color tomarán los valores dados por las expresiones (5-8) y (5-9):

$$Cb = 0.564 (B - Y). \quad (5-8)$$

$$Cr = 0.713 (R - Y). \quad (5-9)$$

El margen dinámico de RGB es de 0 a 0,7 V (a las que se les añaden las señales de sincronismo para completar el valor de 1 V pico a pico). La señal de luminancia varía entre 0 y 700 mV y la de crominancia entre -0,350 a 0,350 V y su nomenclatura en general es Pb y Pr para analógico y Cb y Cr para digital.

5.4.2 Muestreo

La estructura de muestreo se denomina estática cuando se conoce la posición de cada muestra sobre la imagen. Esto implica que existe una relación matemática entre la frecuencia de muestreo y la frecuencia de imagen y por lo tanto también con la frecuencia de línea.

En estas condiciones, las posiciones de las muestras se encuentran perfectamente establecidas y cuando se finalice el muestreo de una imagen, sobre la imagen siguiente volverán a situarse las muestras en las mismas posiciones que en la imagen anterior.

Dentro de la estructura de muestreo estática, se pueden establecer varios tipos de muestreo: ortogonal e intercalado. El utilizado es el muestreo ortogonal.

Muestreo estático ortogonal

Este tipo de muestreo genera unas muestras que sobre una línea en concreto se va a producir la misma situación para todas las imágenes. Dicho de otra manera, con este tipo de muestreo para todas las imágenes se muestrea los mismos puntos, tal y como se aprecia en la Figura 5-9.

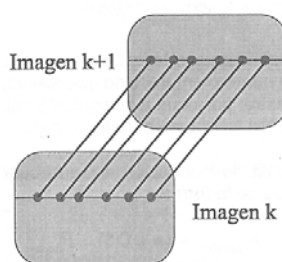


Figura 5-9. Muestreo ortogonal estático.

Para que se produzca esta situación, en cada línea de imagen se han de generar un número entero de muestras, consiguiendo que la posición de muestra sobre cada línea coincida en todas ellas. Es decir, la relación entre la frecuencia de muestreo y la frecuencia de línea deberá cumplir la expresión (5-10):

$$f_{muestreo} = n \cdot f_{línea} \quad (5-10)$$

siendo n un número entero.

Frecuencia de muestreo utilizada

Se sabe que existen dos frecuencias de línea muy utilizadas, la americana (15734,2657 Hz) y la europea (15625 Hz). Para poder establecer una frecuencia de muestreo que sea compatible con ambos sistemas es necesario localizar el mínimo común múltiplo de esas dos frecuencias de línea.

$$15734,2657 \times 143 = 2,25 \text{ MHz}$$

$$15625 \times 144 = 2,25 \text{ MHz}$$

Luego la frecuencia de muestreo debería ser múltiplo de 2,25 MHz. Las posibilidades que se contemplan se muestran en la Tabla 5-2:

Tabla 5-2. Múltiplos posibles de frecuencia de muestreo en vídeo.

Múltiplo n	f [MHz]	Comentarios
1	2,25	Insuficiente para muestrear luma o croma.
2	4,50	Insuficiente para muestrear luma o croma.
3	6,75	No válida para luma, pero es la usada para croma en formato 4:3 y 16:9 normal.
4	9,00	Usada para la señal de croma en formato 16:9 extendido.
5	11,25	No usada.
6	13,50	Usada para la luma en formato 4:3 y 16:9.
7	15,75	No usada.
8	18,00	Usada para la luma en formato 16:9 extendido.

Generalmente para el formato de imagen 4:3 y el 16:9 normal, se usan las frecuencias de muestreo de 13,5 MHz para la señal de luminancia y de 6,75 MHz para la crominancia, siendo esta última la mitad de la primera.

Por otra parte, para el formato de imagen de 16:9 también se contempla la posibilidad de usar las frecuencias de muestreo de 18 y 9 MHz. Para muestrear las señales de luminancia y crominancia respectivamente. En este caso, la densidad lineal de muestras se mantiene respecto a la utilizada para el formato de imagen 4:3, puesto que se cumple que:

$$\frac{16/9}{4/3} = \frac{18}{13,5}$$

Dependiendo del formato de muestreo se usarán distintas frecuencias de muestreo para la luma y la croma, como en el caso de 4:2:2 (el más utilizado), donde la frecuencia de la luminancia es el doble que la de crominancia, o las mismas frecuencias de muestreo para todas las señales, como en el caso del formato 4:4:4. Los formatos posibles de muestreo aparecen en la Tabla 5-3

Tabla 5-3. Formatos de muestreo en televisión digital.

Familia de muestreo	Frecuencia de muestreo para diferentes formatos (MHz)			
	Y	C _B	C _R	Crominancia
4:2:2	13,5	6,75	6,75	En todas las líneas
4:4:4	13,5	13,5	13,5	En todas las líneas
4:1:1	13,5	3,375	3,375	En todas las líneas
4:2:0	13,5	6,75	6,75	En líneas impares
4:1:0	13,5	3,375	3,375	En líneas impares
2:1:1	6,75	3,375	3,375	En todas las líneas

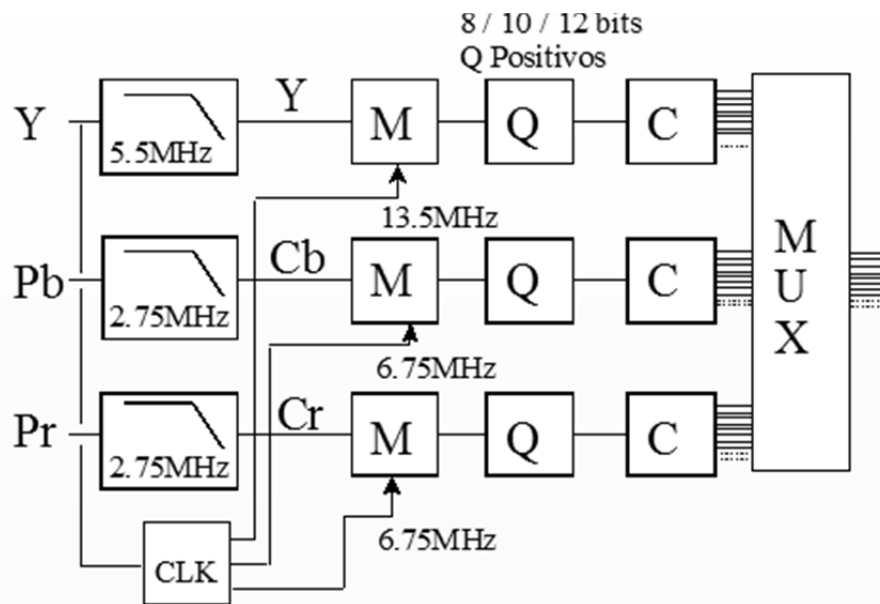


Figura 5-10. Proceso de digitalización de señales analógicas y generación de señal digital en paralelo a 27 Muestras/s.

El proceso de digitalización para cada canal sigue las etapas de muestreo, cuantificación y codificación según el diagrama de bloques de la Figura 5-10. Este proceso que se encuentra triplicado, uno para cada una de las tres señales que van a ser digitalizadas. Según establece el Teorema de Nyquist, la frecuencia de muestreo siempre debe ser superior al doble de la máxima frecuencia que se pretende representar en la señal analógica a digitalizar. Si no se respeta este principio, se provoca aliasing, con el consiguiente deterioro de la calidad de la imagen.

En la Figura 5-11 se representa el filtro anti-aliasing que se utiliza en este caso, de 5.75 MHz para la señal de luminancia, inferior a la mitad de la frecuencia de muestreo de 13,5 MHz.

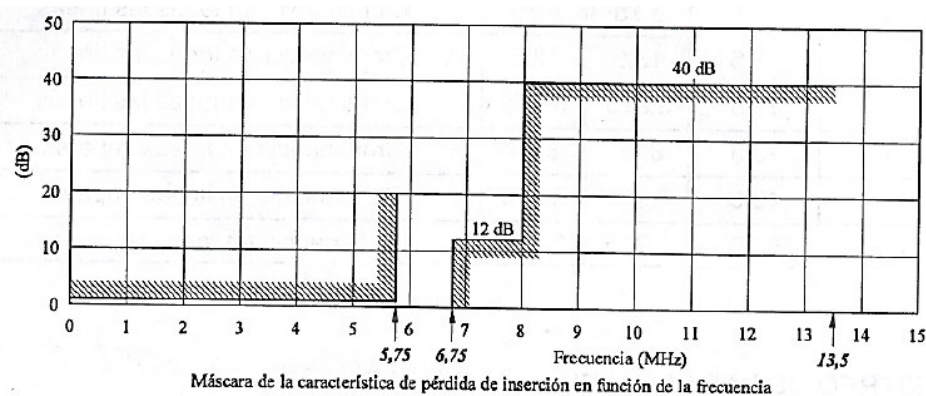


Figura 5-11. Filtro anti-aliasing para la señal de luminancia.

En el caso de las señales de crominancia, Cb y Cr, la frecuencia de corte del filtro anti-aliasing es de 2.75 MHz tal y como se parecía en la Figura 5-12.

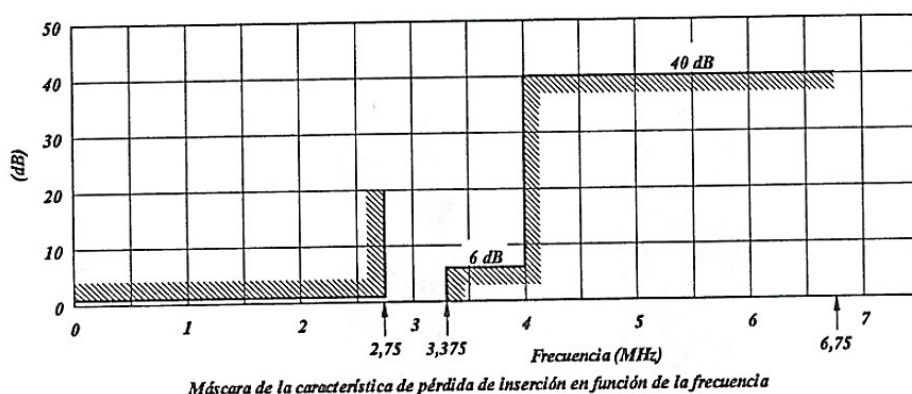


Figura 5-12. Filtro anti-aliasing para la señal de crominancia.

Para cumplir con la plantilla de distorsión de retardo de grupo o distorsión de fase, requieren incorporar un filtro de corrección de fase. Las plantillas permiten un cierto grado de aliasing, pues permiten una atenuación de 12 dB a la frecuencia de Nyquist, lo que no es muy estricto. En vídeo, un ligero aliasing no siempre es detectable, e incluso a veces favorece la sensación psicovisual de “imagen con detalles” (existen estudios estadísticos que confirman esto).

5.4.3 Cuantificación

A la salida del circuito de muestreo se encuentra una señal discretizada en tiempo, pero con valores de amplitudes todavía analógicas, comprendidas entre 0 y 0,7 V para la señal de luminancia y entre -0,350 y +0,350 V para cada una de las dos señales de crominancia.

En la Figura 5-13 se muestra el resultado del conjunto cuantificador-decuantificador, pudiéndose observar que los valores de salida se acercan a los de entrada pero nunca podrán igualarlos, ya que falta información para ello.

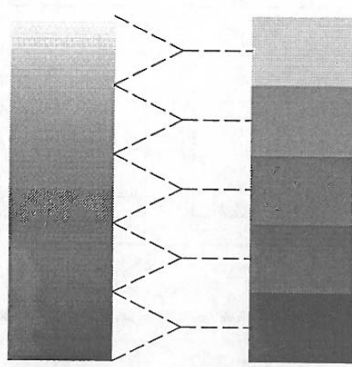


Figura 5-13. Cuantificador-decuantificador.

El valor de salida será el valor central de la franja, por lo que el error cometido en el proceso nunca superará por exceso o por defecto a la mitad del valor del escalón de cuantificación. Luego, aunque el error sea inevitable, se puede controlar y disminuir, para ello hay que controlar el tamaño del escalón de cuantificación y el número de bits que se quiere usar, ya que mientras más pequeño sea el escalón, mayor será el número de franjas y serán necesarios más bits. Como ya se explicó anteriormente en la digitalización de sonido.

Valoraciones empíricas han establecido que con 8 bits (256 niveles) es suficiente para que el ojo humano considere inapreciable la degradación. Posteriormente en una actualización de la Norma se ha considerado la posibilidad de trabajar con 10 bits, para que se parta de una señal con una calidad muy alta que permita bastantes manipulaciones y redondeos posteriores sin que la degradación pase a ser importante.

Con 10 bits el número de niveles es 1024. La cuantificación es lineal y sólo se usan valores positivos. Luego para las señales de crominancia Cb y Cr equivale a desplazar digitalmente 350 mV (offset).

La correspondencia entre niveles analógicos y digitales es la siguiente:

- Y: 877 niveles: Negro 0 mV en 64_d, Blanco 700 mV en 940_d. Resto sobreniveles y reservados sincronismos.
- Cb y Cr: 897 niveles. Sin color 0 mV 512_d. Máximo 350 mV en 960_d. Mínimo -350 mV en 64_d. El resto sobre niveles y sincronismo.

Línea activa digital LAD

En vídeo analógico la duración de una línea es de 64 μs pero el tiempo útil de vídeo es de 52 μs porque el resto del tiempo se dedica a sincronización de línea. Por lo tanto, no resultaría lógico digitalizar la parte no útil de la línea, ya que puede ser usada para la inclusión de otras informaciones.

Normalmente se utiliza el periodo de muestreo como patrón de medida ya que muchos parámetros están relacionados con este valor. Si se tiene una frecuencia de muestreo de 13,5 MHz el periodo viene dado por la expresión (5-11):

$$T = \frac{1}{f_m} = \frac{1}{13,5 \cdot 10^6} = 7,407407 \cdot 10^{-8} \approx 74 \text{ ns. (5-11)}$$

Al digitalizar la parte de vídeo de una línea se utilizan 720 muestras, por lo que la duración de lo que se conoce como Línea Activa Digital y se calcula con la expresión (5-11):

$$\text{LAD} = 720 \cdot 7,407407 \cdot 10^{-8} = 53,33333 \text{ μs.}$$

Para los sistemas europeos en los que la duración de línea analógica es de 64 μs (equivalente a 864 unidades T) la posición relativa de la línea activa digital sobre ella es la que se indica en la Figura 5-14.

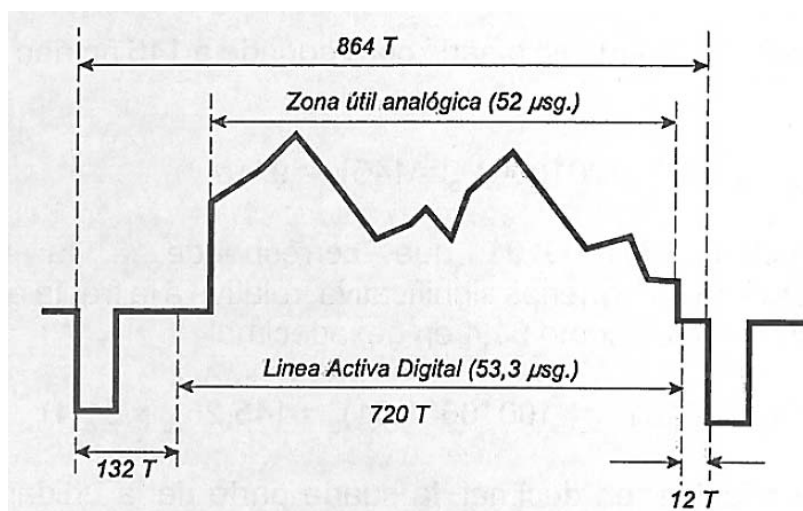


Figura 5-14. Línea activa digital para sistemas de 625 líneas.

Conviene darse cuenta que este ritmo de generar palabras es únicamente durante el tiempo correspondiente a la Línea Activa Digital, porque durante el resto del tiempo no se genera ninguna palabra en absoluto.

5.4.4 Señal de Digital de Vídeo: SDI

Es necesario multiplexar en tiempo las tres señales digitalizadas YCbCr para poder transmitirlos por un único cable o canal. El Serial Digital Interfaz SDI estandariza la forma de transmisión en estudios según ITU-R-BT 656.

Durante el tiempo correspondiente a la LAD, el canal de luminancia genera 13,5 Mmuestras/s y 6,75 Mmuestras/s cada uno de los canales de crominancia. Hay momentos en los que se generan tres palabras simultáneamente y otros en los que solo se genera la muestra de luma.

Estos tres flujos de señal hay que colocarlos en un solo flujo de salida de forma multiplexada según la secuencia que puede apreciarse en la Figura 5-15. Puede comprobarse que es la única forma posible de hacerlo, de manera que se vayan lanzando las palabras al mismo ritmo que se van adquiriendo, sin que sea necesario almacenar información.

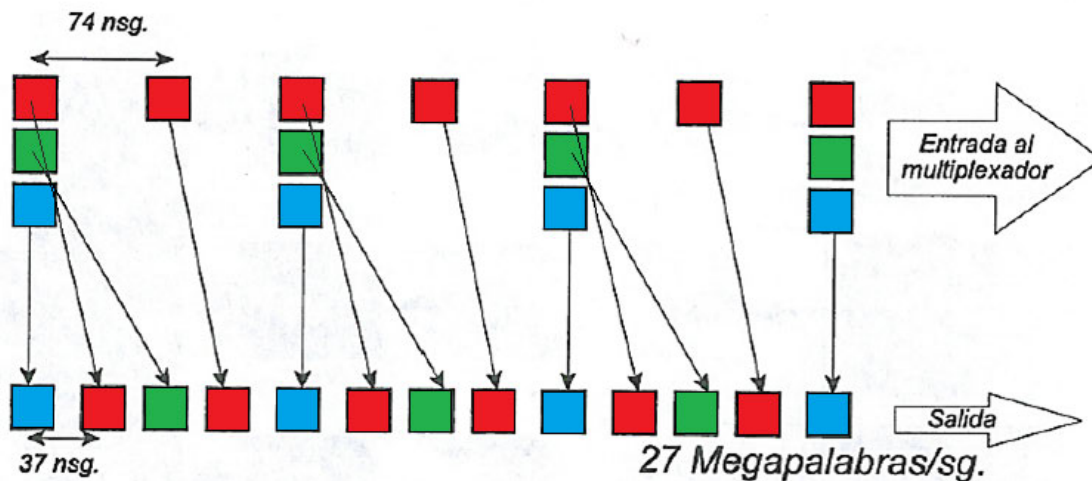


Figura 5-15. Multiplicación de los tres canales de datos.

A la salida del multiplexador se envía una muestra cada 37 ns ($T/2$) de 10. Cuando no hay muestras del LAD se rellena con datos auxiliares.

Cuando finaliza la LAD se envía un cuarteto de palabras (muestras) de sincronismo. De la misma manera, cuando ha finalizado el periodo inactivo deberá incluirse una información que especifique que va a comenzar la siguiente información de vídeo.

Estos cuartetos de palabras reciben los nombres respectivos de EAV (*End Active Video*) y SAV (*Start Active Video*), no solo informan sobre el comienzo y fin de la LAD, también llevan información adicional de sincronismos de líneas y de campos. Su colocación en la secuencia de transmisión es la que se muestra en la Figura 5-16.

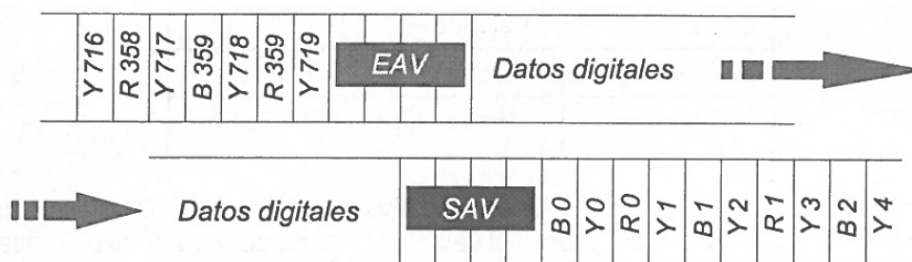


Figura 5-16. Multiplexación de los tres canales de datos.

Puede observarse que en cada una de las líneas queda un total de 280 palabras x 10 bits, que pueden ser utilizadas para introducción de datos, tales como sonido o cualquier otro tipo de informaciones.

Además de la introducción de datos en las partes de la línea que se acaba de comentar, existe también la posibilidad de introducir más datos auxiliares en los intervalos activos de vídeo correspondientes a las líneas de borrado vertical, en las que no se introduce vídeo. Los datos introducidos son canales de audio o metadatos de información de programa.

Cuando no se precisa introducir datos en las zonas asignados a ellos, es necesario rellenar el intervalo correspondiente con palabras de relleno que corresponden al color negro: luminancia y crominancia nulas.

Puede hacerse la transmisión de señales en paralelo, pero lo normal es que se coloque la información en serie, para no necesitar la utilización de conector de multiplex contactos y la correspondiente manguera también de ese

mismo número de cables, con la gran ventaja de que dicha información puede ir soportada en un simple cable coaxial o en una fibra óptica.

Luego el reloj del múltiplex paralelo es de 27 MHz. Otro ejemplos para otros formatos son:

- 27 MHz para 4:2:2 576i 4:3 y 16:9 anamórfico
- 36 MHz para 4:2:2 576i 16:9 mejorado
- 148.5 MHz para 4:2:2 HDTV 720p y 1080i
- 222.75 MHz para 4:4:4 HDTV 720p y 1080i
- 297 MHz para 4:2:2 HDTV 1080p

El circuito serializador, durante el tiempo correspondiente a la LAD, recibe un conjunto de palabras (27 Mmuestras/s), cada una de ellas con 10 bits y debe colocarlas en serie, obteniendo un total de 270 Mbps.

Se coloca en primer lugar en la salida el bit de menor peso de cada palabra y en último lugar el de más peso. El nivel de salida debe estar en 800 mV pico a pico, con una tolerancia de $\pm 10\%$ al ser volcado sobre una carga de 75Ω con un cable de impedancia característica de 75Ω .

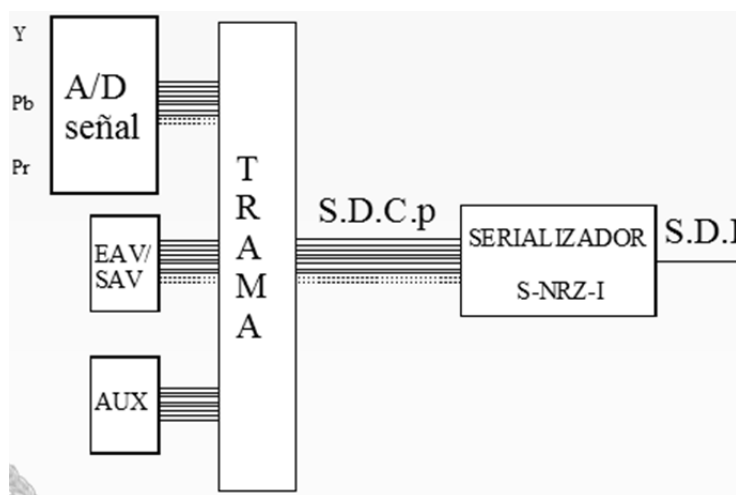


Figura 5-17. Esquema que sigue la señal de vídeo hasta convertirse en SDI.

La señal de vídeo digital se serializa y se le dota de un código de canal y un interface eléctrico adecuados. El código de canal: S-NRZ-I Scrambled-NonReturntoZero-Invertible. Es un código NRZ donde los “0” no producen transición y los “1” producen transición, y donde los bits son aleatorizados antes de codificarse para incrementar el número de transiciones.

Los problemas y límites de la señal SDI son el ancho de banda elevado, la recuperación de reloj, las secuencias patológicas, la respuesta en frecuencia de los cables y las desadaptaciones de impedancias.

5.5 Conversión digital-analógica

En general las normativas de vídeo y audio definen en profundidad los sistemas de digitalización, codificación, transmisión para que todos los equipos receptores puedan ser capaces de utilizar dichas señales. Sin embargo, en el lado de la conversión a analógico y decodificación no existen estándares, y por lo tanto la calidad y diferenciación entre de los equipos receptores dependerá de esta parte del sistema de comunicación, de cómo implemente la decodificación y conversión.

En proceso de conversión digital a analógico (DAC) consiste en que proporcionar un nivel de tensión cuantificado por cada palabra digital binaria indica un número entero. El DAC genera un nivel de tensión específico para cada código entero, y se mantiene hasta que se recibe la siguiente palabra digital binaria, este

tiempo coincide con el tiempo de muestreo. La señal se parecerá a la señal analógica codificada pero su nivel estará cuantificado. Esa señal de escalones se filtra con el filtro paso bajo rector que deje pasar el ancho de banda de la señal original. Los escalones son de frecuencia de muestreo y con el filtro paso banda desaparecerán.

Luego la calidad de la señal reconstruida se ve afectada mucho por la calidad del DAC y su filtro. Es un error habitual no prestar mucha atención a la etapa D/A, que es especialmente crítica, al ser la que genera la señal analógica que se presenta al usuario.

En vídeo, los filtros de reconstrucción están estandarizados y usan la misma plantilla que los de antialiasing.

En audio el conversor también debe implementar la corrección “ $\sin x/x$ ” o “sinc” que compense el filtrado provocado por el uso de pulsos de cierta anchura en la A/D.

Son evidentes las diferencias entre reproductores de CDs, parte de las cuales se deben al filtro de reconstrucción. También son evidentes las diferencias entre reproductores DVD y entre pantallas de vídeo de la misma resolución y tipo, al mostrar el mismo DVD. Parte de ellas se deben al proceso de generación, no siempre se usan los filtros estandarizados.

5.6 Cuestionario

1- La digitalización de una señal de audio o vídeo:

- a) Es la transformación directa de las variaciones de onda en códigos binarios.
- b) Significa muestrear dichas señales para obtener valores concretos de la señal que podrán ser codificados en un proceso posterior.
- c) Es un proceso por el cual se recupera la señal analógica a la entrada del sistema de codificación.-
- d) Ninguna de las opciones anteriores son ciertas.

2-El teorema de Nyquist establece:

- a) La frecuencia máxima a la que se puede muestrear una señal para que no se produzca *overflow* a la salida del digitalizador.
- b) La frecuencia mínima a la que se puede muestrear una señal para que no se produzca *overflow* a la salida del digitalizador.
- c) Establece la frecuencia máxima de muestreo para que no se produzca *aliasing*.
- d) Establece la frecuencia mínima de muestreo para que no se produzca *aliasing*.

3-Para muestrear una señal se multiplica el valor de dicha señal por un tren de deltas con una frecuencia máxima establecida por el teorema de Nyquist.

- a) Verdadero.
- b) Falso.

4-En el proceso de cuantificación:

- a) Se otorga a cada muestra de la señal un valor binario determinado.
- b) Se le añade un error a la señal.
- c) La señal se convierte de nuevo en una señal analógica.
- d) Ninguna de las anteriores es cierta.

5- Calcula la relación S/E_{dB} de una señal que será codificada con palabras de 8 bits. (51,36 dB).

6-La amplitud de los escalones de cuantificación es inversamente proporcional al número de bits de la palabra digitalizada.

- a) Verdadero.
- b) Falso.

7- Con la cuantificación no uniforme se usan diferentes amplitudes de escalones de cuantificación, se eligen pequeños escalones allí donde sea más probable la variable de entrada y escalones grandes para amplitudes menos probables.

- a) Verdadero.
- b) Falso.

8- Calcular para la ley A la mejora que se produce en la SRN_Q para una línea de telefonía estándar. (87,6 dB)

9- El Dither es una técnica digital usada para eliminar los problemas causados por el error de cuantificación en señales de bajo y alto nivel.

- a) Verdadero.
- b) Falso.

10-Con el Dither:

- a) No se enmascara el error de cuantificación.
- b) Se elimina el error de la señal de audio.
- c) Se lleva el error a bajas frecuencias, de forma que no se puede percibir.
- d) Ninguna de las anteriores es cierta.

11- En vídeo se digitaliza la señal en diferencias de color porque:

- a) Su tratamiento es más simple debido a que su rango de amplitudes es mínimo.
- b) El ojo humano es muy sensible a la señal de luminancia y menos exigente con las señales de crominancia.
- c) Porque cada una de las tres señales, la luminancia y las dos señales de crominancia, llevan la información correspondiente a los colores rojo, azul y verde.
- d) Ninguna de las opciones anteriores es cierta.

12- El muestreo estático ortogonal es un tipo de muestreo que genera unas muestras que sobre una línea en concreto se va a producir la misma situación para todas las imágenes.

- a) Verdadero.
- b) Falso.

13- El muestreo ortogonal intercalado es un tipo de muestreo que genera unas muestras que sobre una línea en concreto se va a producir la misma situación para todas las imágenes.

- a) Verdadero.
- b) Falso.

14- Asocia cada frecuencia de muestreo a su descripción.

- | | |
|---------------|--|
| a) 6,75 Mhz. | 1- No usada. |
| b) 9,00 Mhz. | 2- Usada en la luma en formato 16:9 extendido. |
| c) 11,25 Mhz. | 3- Usada para la croma en formato 16:9 extendido. |
| d) 13,50 Mhz. | 4- Usada para la luma en los formatos 4:3 y 16:9. |
| e) 18,00 Mhz. | 5- No es válida para luma, pero se usa para croma en formatos 4:3 y 16:9 normal. |

15-Con el formato de muestreo 4:2:0 la crominancia se muestrea en todas las líneas.

- a) Verdadero.
- b) Falso.

16-La línea digital activa:

- a) Dura lo mismo que el tiempo útil de vídeo analógico.
- b) Es la línea sustituta de la EAV.
- c) Necesita recibir la información directamente de la línea activa digital de la señal analógica.
- d) Ninguna de las anteriores es cierta.

17- El orden es:

- a) Muestrear, cuantificar y codificar.
- b) Muestrear, codificar y cuantificar.
- c) Cuantificar, codificar y muestrear.
- d) Codificar, cuantificar y muestrear.

18-EAV es:

- a) La señal que indica que una secuencia de vídeo termina.
- b) La señal que indica el inicio de vídeo activo.
- c) Indica el fin de vídeo activo.
- d) Ninguna de las anteriores es cierta.

19-SAV es:

- a) La señal que indica que una secuencia de vídeo termina.
- b) La señal que indica el inicio de vídeo activo.
- c) Indica el fin de vídeo activo.
- d) Ninguna de las anteriores es cierta.

20-La señal de vídeo se serializa para su transmisión.

- a) Verdadero.
- b) Falso.

6 CODIFICACIÓN DE LA SEÑAL DE AUDIO

La codificación es el proceso por el cual se obtiene una representación digital de la señal analógica. La digitalización PCM es la codificación básica y da lugar a un régimen binario básico.

De forma general la codificación puede tener dos objetivos. El primero es disminuir el régimen binario de la señal codificada basándose en diferentes tipos de codificación. A este tipo de codificadores se les llama codificadores fuente. El segundo es adaptar la señal al medio de transmisión, a este tipo de codificadores se les llama codificadores de canal.

Las codificaciones pueden ser sin pérdidas (Lossless), con pérdidas no perceptibles (Lossy) o con pérdidas perceptibles que darán regímenes binarios cada vez más pequeños.

De las diversas técnicas para codificar audio las más extendidas y de mayor éxito son las denominadas codificaciones perceptuales, cuyas decisiones están basadas en el comportamiento psicoacústico del sistema auditivo humano. En concreto se basa en el fenómeno de **enmascaramiento auditivo**. Estas codificaciones pueden ajustarse para ser bien Lossy o bien con pérdidas, dependiendo de cuan bajo se quiera el régimen binario resultante.

Empleando un modelo psicoacústico del sistema auditivo humano, el codificador identifica el contenido imperceptible de la señal y no lo codifica, produciendo un flujo final de bits que representa sólo la información perceptible.

6.1 Psicoacústica

La psicoacústica estudia la percepción auditiva humana, desde el diseño biológico del oído hasta la interpretación cerebral de la información auditiva.

El oído, y todo el sistema nervioso asociado, son enormemente complejos e interactivos; pero pese a su complejidad, también cuenta con una serie de limitaciones. Por una parte, el oído es muy agudo en detectar ciertos matices o defectos de la señal, pero es muy torpe con otros defectos, considerándolos prácticamente irrelevantes.

La respuesta del oído con la frecuencia es logarítmica. Por ejemplo, el intervalo entre 100 y 200 Hz se percibe como una octava, igual que el intervalo entre 1 kHz y 2 kHz. En términos lineales, la segunda octava es mucho más grande, pero sin embargo el oído lo percibe como si se tratara de un intervalo de variación de frecuencia idéntico.

La frecuencia es una medida física, sin embargo el tono es una característica subjetiva y compleja basada en la frecuencia y en otros parámetros físicos como la forma de onda y la intensidad. Por ejemplo, si se reproduce una senoide de 200 Hz con dos niveles distintos, la mayoría de los oyentes dirán que la senoide más fuerte tiene un tono menor. De hecho, puede llegar a ser necesario incrementar un 10% la frecuencia para preservar la percepción subjetiva del tono (en frecuencias bajas). Por otra parte, en la parte más sensible del oído, de 1 a 5 kHz, no existe casi variación del tono con la intensidad.

El batido de frecuencia es un fenómeno que aparece cuando se escuchan simultáneamente dos frecuencias cercanas. Cuando la diferencia entre ambas frecuencias tiene un valor que corresponde a una frecuencia audible, este tono diferencia se puede llegar a percibir de forma nítida. Este efecto es especialmente audible cuando las frecuencias son altas y los tonos tienen una intensidad elevada.

El oído es capaz de recibir un gran margen dinámico de niveles; el umbral de dolor a 120 dB_{SPL} tiene una intensidad sonora de 1.000.000.000 de veces el umbral de audición a 0 dB_{SPL}. Por eso se usa una escala logarítmica y no una lineal para representar dicho margen dinámico.

La sensibilidad del oído depende de la frecuencia. La máxima sensibilidad se encuentra entre 1 y 5 kHz, con pérdida de sensibilidad en frecuencias altas y bajas. En la Figura 6-1 se muestran las curvas de igual sonoridad, obtenidas con pruebas de escucha sobre un gran número de sujetos. Cada curva representa el margen de frecuencias percibidas con una misma sonoridad. La curva inferior corresponde al umbral de audición: el mínimo nivel de presión sonora en función de la frecuencia que puede percibir una persona con un oído normal. Por ejemplo, un tono de 30 Hz apenas audible, tiene una intensidad 60 dB mayor que un tono de 4 kHz de un nivel equivalente al umbral de audición. La respuesta varía en función de la intensidad; cuanto más intenso sea el sonido, más plana es la respuesta. Las curvas se expresan en fonios, nivel de presión sonora (SPL) de cada curva en la frecuencia de 1 kHz.

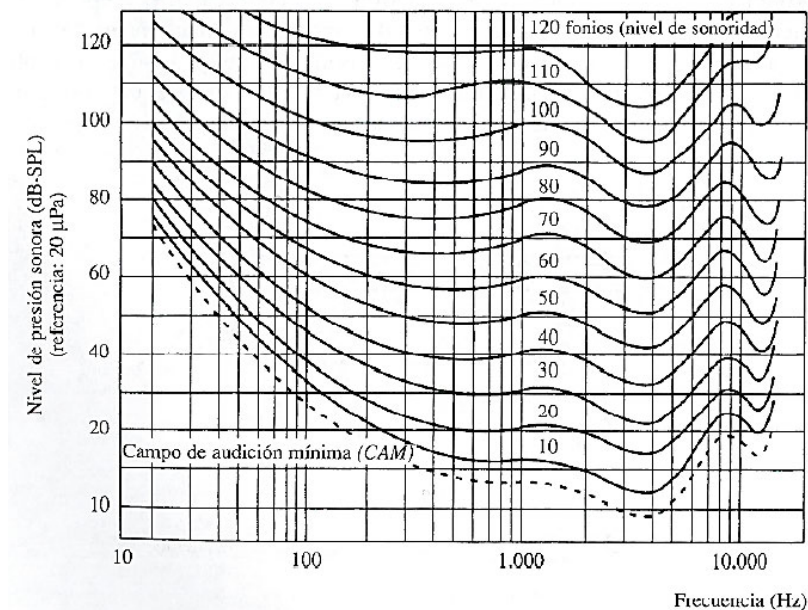


Figura 6-1. Las curvas de igual sonoridad Robinson-Dadson muestran la no linealidad del oído con la frecuencia y con el nivel. Estas curvas están basadas en estudios psicoacústicos, utilizando sinusoides como señales de prueba.

El oído percibe solo una parte de toda la información que contiene la señal de audio; esta información se llama entropía perceptual. Las señales con poca entropía pueden ser codificadas perceptualmente de forma muy eficiente (gran reducción binaria). Por esta razón, un codificador de audio deberá diseñarse con una tasa binaria de datos variable: baja cuando la información sea pobre y alta cuando haya más información.

La salida de datos debe ser variable porque aunque la frecuencia de muestreo sea constante la entropía de la señal no lo es. Aplicando técnicas psicoacústicas es posible eliminar las partes irrelevantes de la señal; a este proceso se le denomina reducción de datos. La señal original no se puede reconstruir de forma exacta. La función de cualquier sistema de reducción de datos es eliminar la entropía de la señal, aplicando modelos perceptuales de entropía es posible eliminar las partes irrelevantes de la señal, de aquí que la reducción de datos obtenida sea inaudible. Un codificador perceptual debe aplicar modelos psicoacústicos para identificar el contenido irrelevante de las señales de audio.

Las ventajas que ofrece la codificación perceptual, gracias la reducción del régimen binario son:

- La reducción de los costes de los canales de transmisión.
- La optimización de la capacidad de almacenamiento de un soporte.
- La optimización de los canales de transmisión para varios servicios simultáneos (datos, imágenes, voz, audio, internet, etc.).

Se puede aplicar codificación perceptual a la transmisión, almacenamiento, videoconferencia, telefonía, multimedia interactiva, etc. Pero su uso conlleva una serie de requerimientos tales como la alta calidad de las señales decodificadas, decodificadores robustos que soporten bien las variaciones de nivel y de espectro,

robustez también frente a errores (ha de incluir sistemas correctores), decodificadores simples y baratos, generar los mínimos retardos posibles, posibilidad de mezclar, editar, atenuar y comprimir el rango dinámico de diferentes señales codificadas y una correcta sincronización de audio y vídeo.

6.2 Fisiología del oído humano y bandas críticas. Percepción auditiva

Simplificando el comportamiento del oído interno, este se comporta como un banco de filtros de ancho de banda variable llamadas **bandas críticas** (aproximadamente de tercio de octava). De forma empírica se ha demostrado que las bandas críticas son mucho más estrechas en frecuencias bajas que en frecuencias altas; las tres cuartas partes de las bandas críticas se encuentran por debajo de los 5 kHz; lo cual se traduce en que en graves podemos distinguir mejor frecuencias cercanas entre sí (de unos pocos hercios) mientras que en alta frecuencia hacen falta centenares de hercios para que podamos distinguir una frecuencia de otra.

6.2.1 Umbral de audición

Los dos fenómenos que gobiernan la audición humana son el umbral de audición y el nivel de enmascaramiento, tal como se muestra en la Figura 6-2. La curva del umbral de audición representa el mínimo nivel (0 sonios) que necesita el oído para detectar un tono de una frecuencia dada. El umbral está referenciado a 0 dB_{SPL} en la frecuencia de 1 kHz. Un sonio representa la sonoridad de un tono de 1 kHz con un nivel de 40 dB_{SPL}. Una sonoridad de 2 sonios corresponde a un nivel de 50 dB_{SPL}; de forma similar, cada vez que se duplica la sonoridad (en sonios) se produce un incremento de 10 dB en el nivel de presión sonora. Por ejemplo, 64 sonios corresponden a un nivel de 100 dB_{SPL}. En un codificador perceptual la señal de entrada se compara con el umbral mínimo, descartándose cualquier señal cuyo nivel esté por debajo de éste, ya que el oído será incapaz de percibirla.

6.2.2 Enmascaramiento frecuencial

El enmascaramiento frecuencial ocurre cuando un tono desplaza la curva umbral en un margen de frecuencias circundante (frecuencias más altas). El umbral de enmascaramiento representa el mínimo nivel de un tono para que éste sea audible en la presencia de otro. Cuando existen varios tonos simultáneos, los de nivel alto pueden enmascarar completamente a los tonos de nivel inferior y cercanos en frecuencia. Por ejemplo, un tono de 500 Hz puede enmascarar a un tono de 600 Hz de menor intensidad. En otras palabras, la existencia de un sonido ciertamente no asegura su audibilidad y a la inversa, puede llegar a asegurar la audibilidad de otro. Al sonido más fuerte se le denomina sonido enmascarante y al más débil sonido enmascarado. La teoría del enmascaramiento dice que el tono más débil sólo es detectado cuando su energía se iguala a la energía contenida en la banda crítica de la señal enmascarante; esta relación es lineal con respecto a la amplitud.

La membrana basilar es responsable del enmascaramiento en amplitud. Una respuesta intensa en un punto de la misma enmascarará respuestas más débiles que se produzcan dentro de la banda crítica circundante. Esto se debe a que hasta que el sonido llegue a la zona de la membrana que se encarga de procesarlo antes se han excitado otras células de bandas críticas de frecuencias más altas. La distancia entre el nivel de la señal enmascaradora y el umbral de enmascaramiento se le denomina relación Señal/Enmascaramiento (SMR).

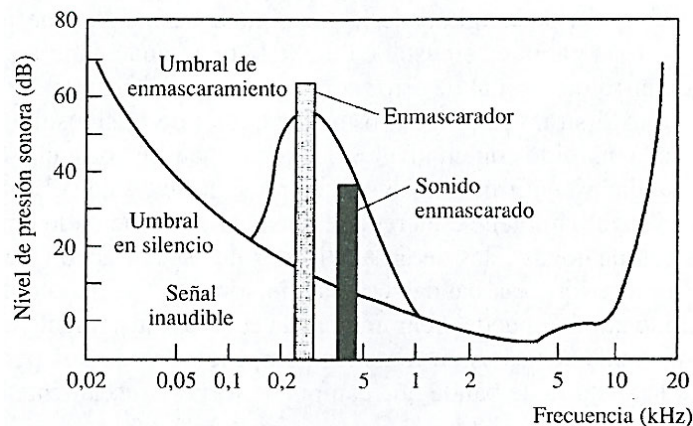


Figura 6-2. El umbral de audición describe los sonidos más débiles que son audibles a lo largo del margen de audición humana. Un tono o ruido enmascarador aumenta el umbral de audición en una cierta región, creando una nueva curva de enmascaramiento. Los tonos o ruidos enmascarados, tal vez audibles en otras circunstancias, que están por debajo de la curva de enmascaramiento durante ese intervalo de tiempo no serán audibles.

El ruido de cuantificación será audible siempre que dicho nivel sea inferior al umbral de enmascaramiento: $SNR > SMR$.

La codificación perceptual se centra en conocer como la señal enmascara al ruido de cuantificación. Las bandas de sonido producen curvas de enmascaramiento que son distintas a las producidas por señales sinusoidales; normalmente, la música tiene unas propiedades enmascarantes muy parecidas a las de los tonos. En los modelos de enmascaramiento se utilizan curvas de enmascaramiento de señales sinusoidales, ya que éstas aseguran el menor nivel de enmascaramiento (peor caso) sobre el ruido de cuantificación; las señales tonales algo más complejas producen un mayor enmascaramiento. Generalmente, un tono es inaudible cuando su nivel es aproximadamente 4 dB por debajo de la curva de enmascaramiento que produce la señal de una banda crítica. De forma contraria, para que una banda de 1/3 de octava sea enmascarada por un tono puro, el nivel debe ser 24 dB inferior al del tono; por tanto, existen 20 dB de diferencia. Al final, el ajuste fino de enmascaramiento debe hacerse después de haber escuchado la señal codificada. Además, todos los codificadores tienen unos umbrales por debajo de las curvas estándar, ya que éstos no pueden saber a priori cuál será la calidad de la señal codificada. Apenas existen estudios de enmascaramiento realizados directamente con señales de música. Sin embargo, se puede considerar que los modelos de enmascaramiento con tonos son válidos para la señal musical. Se debe continuar investigando más sobre el enmascaramiento producido por señales musicales.

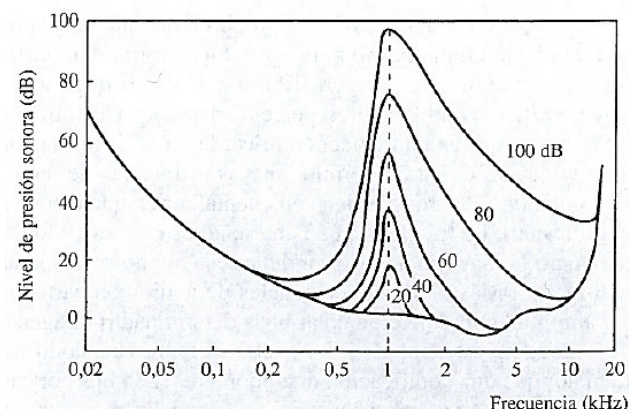


Figura 6-3. Los umbrales de enmascaramiento varían en función del nivel de presión sonora; en esta prueba se ha utilizado un ruido de banda estrecha centrado a 1 kHz. La curva a bajas frecuencias queda prácticamente inalterada.

6.2.3 Enmascaramiento temporal

El enmascaramiento temporal aparece cuando los tonos se suceden rápidamente, pero no de forma simultánea. Una señal puede llegar a ser enmascarada por otra señal que aparezca con posterioridad; a este fenómeno se le denomina pre-enmascaramiento (algunas veces llamado enmascaramiento regresivo). Además, una señal puede ser enmascarada por otra señal que se haya extinguido. A este fenómeno se le denomina post-enmascaramiento (o enmascaramiento regresivo). El enmascaramiento temporal es tanto mayor cuanto más próximos estén en el tiempo la señal enmascarante y la señal enmascarada, de la misma forma que un enmascaramiento frecuencial es tanto mayor cuanto más próximas sean las frecuencias de ambas señales: enmascarante y enmascaradora. Dado un tono de 80 dB_{SPL}, puede existir un post-enmascaramiento de 40 dB dentro de un intervalo de 20 ms y de 0 dB a partir de 200 ms. El pre-enmascaramiento puede proporcionar hasta 60 dB de enmascaramiento a 1 ms y de 0 dB a 25 ms. Esto se muestra en la Figura 6-4. El enmascaramiento temporal decrece a medida que disminuye la duración de la señal enmascarante. Además, un tono se enmascara por un tono precedente cuando están próximos en frecuencia o cuando este último es de una frecuencia inferior; el post-enmascaramiento es menor cuando el tono enmascarante es de mayor frecuencia. Lógicamente, el enmascaramiento simultáneo es mucho más fuerte que el post y que el pre-enmascaramiento, ya que los sonidos ocurren de forma simultánea.

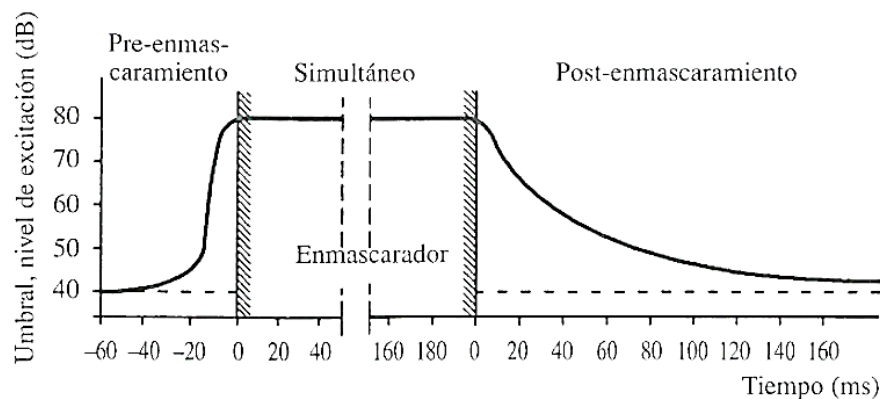


Figura 6-4. El enmascaramiento temporal ocurre antes, y en particular, después de un sonido enmascarador; el umbral decrece con el tiempo. La línea discontinua representa el umbral para tonos impulsivos en ausencia de señal enmascaradora (Zwicker).

En el enmascaramiento temporal, que se muestra en la Figura 6-5 B, se supone que el cerebro integra el sonido en pequeños intervalos de tiempo y los procesa en la corteza cerebral. Además, el cerebro suele procesar más rápidamente los sonidos fuertes que los débiles. En cualquiera de los dos casos, el enmascaramiento temporal es de suma importancia en una codificación perceptual.

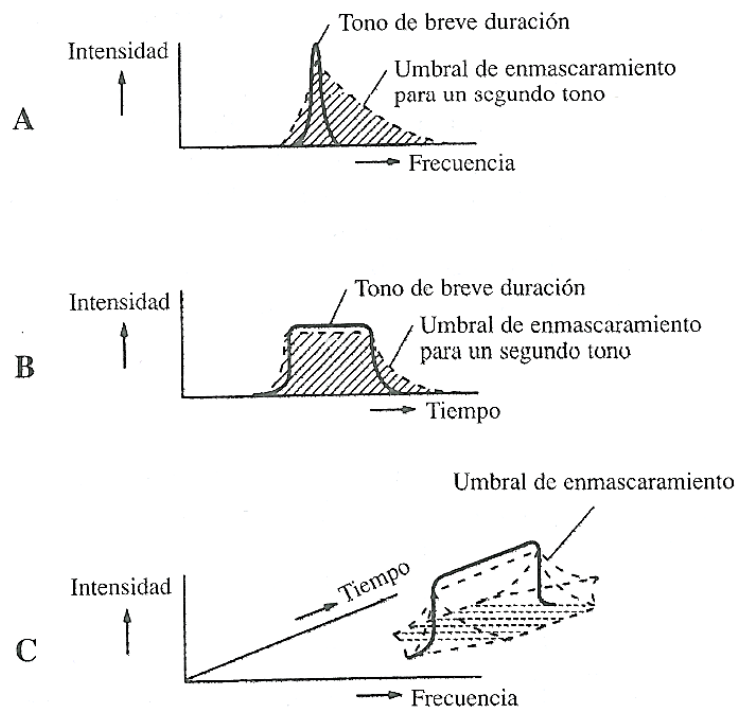


Figura 6-5. Contorno tiempo-frecuencia resultado de un enmascaramiento simultáneo en tiempo y en frecuencia. A. Un codificador perceptual debe colocar un ruido de cuantificación y otros errores de codificación por debajo de este contorno con objeto de hacerlos inaudibles. B. Enmascaramiento temporal. C. Enmascaramiento combinado tiempo-frecuencia.

Los codificadores tienen una limitada resolución temporal, ya que trabajan con bloques de muestras, extendiendo el error de codificación uniformemente en todo un intervalo de tiempo; el enmascaramiento temporal puede hacer inaudibles los chasquidos esporádicos (*artifacts*) producidos por señales transitorias. Idealmente, los bancos de filtros deben proporcionar una resolución temporal entre 2 a 4 ms. El enmascaramiento conjunto (frecuencial y temporal) forma una 'campana' espectral enmascarante, tal como se muestra en la Figura 6-5 C. Los sonidos que estén por debajo de la misma serán enmascarados; los codificadores perceptuales deben identificar las variaciones dinámicas que sufre esta 'campana' para poder codificar la señal de forma correcta.

6.3 Fundamentos de la codificación perceptual

El objetivo de un sistema de reducción de datos es disminuir la tasa binaria: el producto de la frecuencia de muestreo por la longitud binaria de los datos. Esto se puede conseguir fácilmente si se reduce la frecuencia de muestreo pero se produciría aliasing en la señal; o la longitud de la palabra binaria pero se reduciría el margen dinámico (6 dB por cada bit que se suprime).

Generalmente la filosofía de la codificación era generar a la salida del codificador una onda lo más parecida posible a la onda de entrada, luego son importantes todos y cada uno de los datos que la señal de entrada llevaba consigo, incluso los que no somos capaces de percibir. Pero para la codificación perceptual es más importante la calidad de percepción (comprensión) por parte del oyente que preservar la forma de onda de la señal.

Los codificadores perceptuales analizan el contenido de la señal, tanto en nivel como en frecuencia, y lo comparan con un modelo psicoacústico de percepción auditiva. Con este modelo, el codificador elimina la irrelevancia y la redundancia estadística de la señal de audio. En teoría, aunque este modelo produzca pérdidas de componentes en la señal, el sistema auditivo es incapaz de percibir tal degradación de la señal. Se pueden llegar a alcanzar altas tasas de reducción binaria. Por ejemplo, un codificador perceptual puede reducir el régimen binario de un canal de datos de 768 kb/s a 128 kb/s; los datos de 16 bits se reducen a un promedio de 2,667 bits, lo que supone una reducción binaria superior al 83%. En la Tabla 6-1 se reflejan varias tasas de

reducción y los regímenes binarios resultantes para señales monocal y para frecuencias de muestreo de 48 kHz y 44,1 kHz. Los codificadores perceptuales son tan eficientes que tan solo requieren una pequeña parte de los datos grabados por el sistema digital convencional.

Tabla 6-1. Reducción de la tasa binaria con frecuencia de muestreo de 48 y 44,1 kHz.

Bits/muestra	Relación de compresión	48 kHz	44,1 kHz
16	1:1	768 kbps	705,6 kbps
8	2:1	384	352,8
4	4:1	192	176
2,67	6:1	128	117,7
2	8:1	96	88,2
1,45	11:1	69,6	64

Mientras que en los codificadores PCM asignan el mismo número de bits a toda la información de la señal de audio, los codificadores perceptuales si registran un nivel alto asignan mayor número de bits a dicho nivel para asegurar su perceptibilidad, mientras que si se encuentran con un nivel bajo el número de bits que se le asigna es menor.

6.3.1 Asignación de bits

Normalmente, en una codificación perceptual se utilizan dos estrategias de asignación de bits.

a) Asignación adaptativa progresiva

Toda la asignación binaria es realizada por el codificador, y la información de asignación se integra en la trama de datos. La asignación de bits debe ser muy precisa, lo que obliga a que el diseño del codificador sea bastante complejo. Una característica importante a resaltar en la asignación adaptativa progresiva es que el modelo psicoacústico se encuentra únicamente en el codificador; el decodificador no necesita ningún modelo psicoacústico, ya que en la reconstrucción de la señal de audio solo utiliza los datos codificados. Una desventaja en este tipo de asignación es que parte del régimen binario codificado se emplea para llevar información colateral (no audio) referente a la asignación binaria aplicada, necesaria en el decodificador para poder reconstruir la señal original.

b) Asignación adaptativa regresiva

La información colateral de asignación binaria se deduce directamente de los datos codificados de audio, sin necesidad de que el codificador envíe ningún tipo de información suplementaria. El régimen binario de la señal codificada no lleva información colateral de asignación. Sin embargo, puesto que el decodificador necesita extraer esta información de los propios datos de audio, este tipo de codificación no llega a ser tan precisa como la anterior. Además, el decodificador es bastante más complejo y el modelo psicoacústico no admite mejoras.

La codificación perceptual es normalmente tolerante a los errores de codificación. En una codificación PCM, el error de codificación se traduce en un ruido de banda ancha; sin embargo, en la mayoría de los codificadores perceptuales el error se limita a un margen estrecho de frecuencia que corresponde al de la banda crítica codificada; de esta forma su percepción auditiva disminuye. Un error de codificación se percibe como un ruido de bajo nivel en forma de salva. Esto y las técnicas de corrección de errores incluidas en los codificadores perceptuales hacen a este sistema muy robusto frente a los errores de codificación.

La codificación perceptual se puede resumir en los siguientes puntos:

La codificación se determina a partir de la curva SMR vs frecuencia.

La relación SMR de cada banda crítica determina el mínimo número de bits necesarios de esa banda crítica para hacer que $SNR > SMR$.

El codificador determina el régimen binario con el que se establece el número de bits disponibles para cada banda en función de su SMR.

Aquellas bandas cuyo SMR sea negativo no se codifica, ya que un $SMR < 0$ quiere decir que la señal es enmascarada por otra y por tanto, inaudible.

Si se consigue mantener $SNR > SMR$ en cada banda de frecuencia no habrá diferencia perceptual entre la señal original y la codificada.

Si el régimen binario fijado no es suficiente, se rebaja el número de bits en cada banda, aunque $SNR < SMR$. En este caso se estaría cometiendo un error: se añade un nivel de ruido espectral con la misma forma que tiene la curva SMR en función de la frecuencia.

6.4 Métodos de codificación

6.4.1 Codificación digital de la forma de onda

El principal objetivo de este sistema es preservar la forma de onda de la señal en el proceso de codificación. Está basada en la auto-correlación de la señal y es un sistema aplicable para codificar señal de voz y en menor medida la señal de audio.

Mediante este método se codifica el error entre la señal y una predicción de la misma. Dado que el error es una señal de nivel muy bajo necesita menos bits para ser codificada. Se conoce como DCPM (*Differential Pulse Code Modulation*) a este método de codificación, cuyo esquema se detalla en la Figura 6-6.

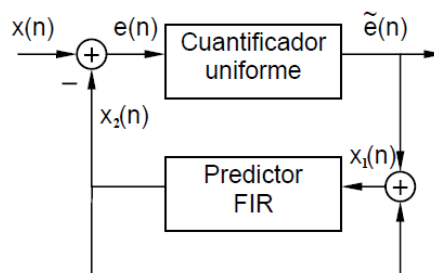


Figura 6-6. Codificador DCPM. $x(n)$: dato de entrada, $e(n)$: error de predicción, $\tilde{e}(n)$: error cuantificado, $x_1(n)$: dato cuantificado, $x_2(n)$: dato estimado.

Con el codificador DCPM se consigue reducir en un bit la longitud de palabra o bien, sin reducción de la longitud de palabra, se ganan 6 dB en la SNR de la señal. Según el esquema del DCPM es la calidad del predictor FIR la que define la calidad del propio codificador.

En la Figura 6-7 se muestra el decodificador DCPM.

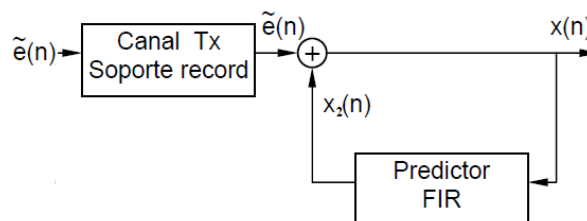


Figura 6-7. Decodificación DCPM. $\tilde{e}(n)$: error cuantificado, $x_2(n)$: dato estimado, $x(n)$: secuencia inicial.

Una mejora del cuantificador DCPM es el ADPCM (*Adaptive Differential Pulse Code Modulation*). En este último se cuenta con cuantificación no uniforme y con un predictor FIR adaptativo. Con ello se consigue reducir un 75% el régimen binario. El esquema de su codificador y decodificador se muestran en la Figura 6-8 y en la Figura 6-9 respectivamente.

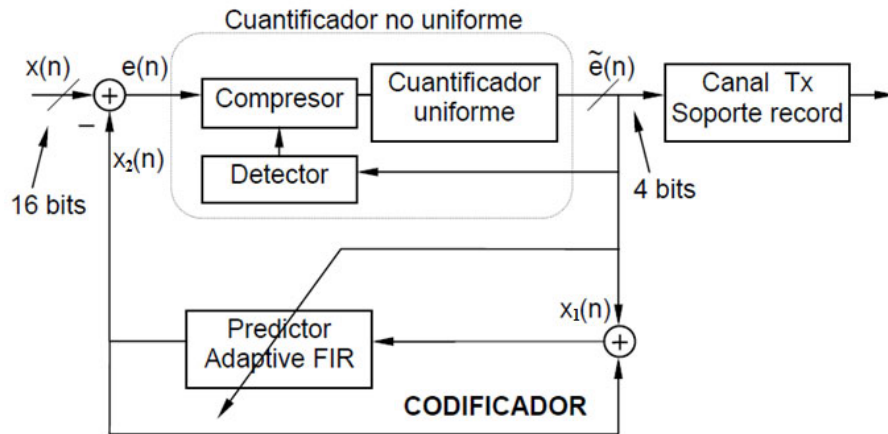


Figura 6-8. $x(n)$: dato de entrada, $e(n)$: error de predicción, $\tilde{e}(n)$: error cuantificado, $x_1(n)$: dato cuantificado, $x_2(n)$: dato estimado.

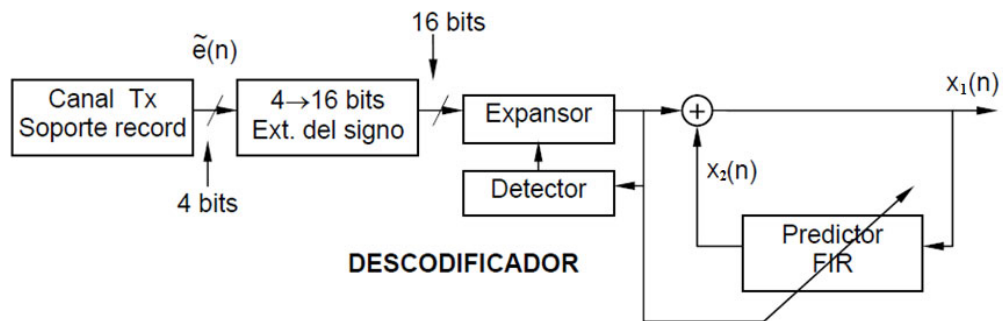


Figura 6-9. $e(n)$: error de predicción, $\tilde{e}(n)$: error cuantificado, $x_1(n)$: dato cuantificado, $x_2(n)$: dato estimado.

6.4.2 Codificación de audio

La codificación de audio está basada en el contenido espectral de la señal, es la frecuencia del sonido lo que lleva información para nuestros oídos. Existen dos formas de codificación de audio:

- Codificación por sub-bandas.
- Codificación adaptativa por transformada.

Normalmente los codificadores sub-banda utilizan un pequeño número de sub-bandas y procesan las muestras directamente en el dominio del tiempo, mientras que los codificadores de transformada utilizan un gran número de sub-bandas y procesan las muestras en el dominio de la frecuencia.

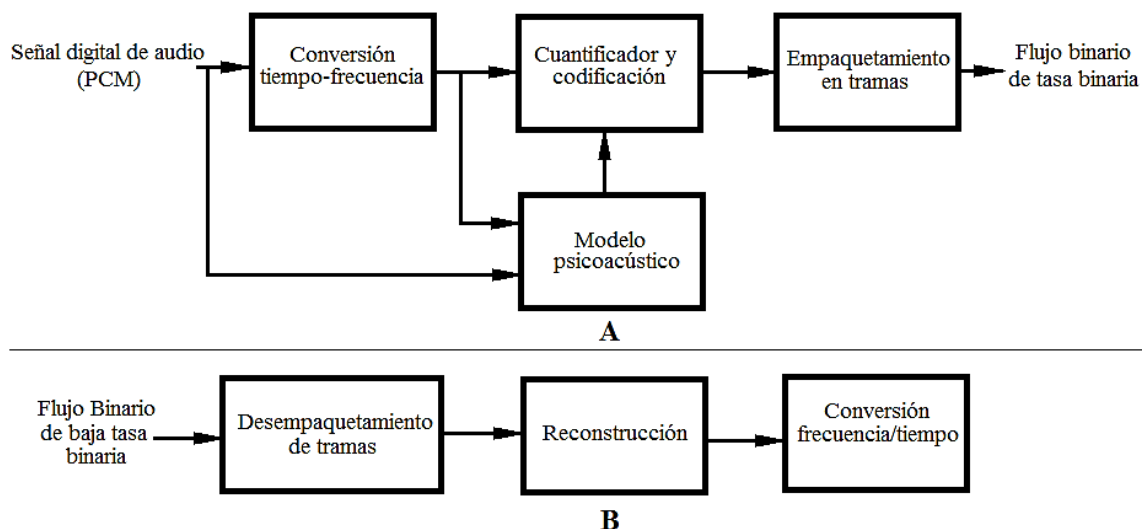


Figura 6-10. Estructura básica de un codificador y un decodificador frecuencia-tiempo (A y B respectivamente). Los codificadores sub-banda cuantifican muestras temporales y los codificadores por transformada cuantifican coeficientes espectrales.

Generalmente, los codificadores sub-banda tienen una buena resolución en tiempo y una mediocre resolución en frecuencia, mientras que en los codificadores por transformada ocurre todo lo contrario. Sin embargo, la diferenciación entre ambos tipos de codificación es más bien académica; matemáticamente, todas las transformadas usadas en los codificadores pueden ser consideradas como bancos de filtros. Ambos tipos de codificadores siguen la arquitectura mostrada en la Figura 6-10; las muestras en el dominio del tiempo, o los coeficientes en el dominio de la frecuencia, son codificados de acuerdo a un modelo psicoacústico presente en el codificador.

La codificación sub-banda es un híbrido entre técnicas que trabajan en el dominio del tiempo y en el dominio de la frecuencia. Un pequeño bloque de muestras de entrada de banda ancha se divide en un determinado número de bandas mediante un banco de filtros paso banda, de esta forma es posible determinar la energía en cada una de las bandas. Mediante un análisis espectral realizado en paralelo con una transformada de Fourier, se analizan las muestras de cada sub-banda para evaluar su energía y codificar la señal de acuerdo con el modelo psicoacústico. Para obtener una gran resolución espectral es preciso trabajar con grandes bloques de datos, que en algunos casos puede llegar a ser un verdadero problema. Los codificadores de baja tasa binaria trabajan directamente con los bloques de muestras. Estos bloques deben ser lo suficientemente pequeños para que estén dentro de la resolución temporal del oído. El ruido de cuantificación se extenderá por toda la banda en el intervalo de duración del bloque. Si el bloque es demasiado grande, el ruido puede llegar a ser audible en forma de molestos **pre-ecos** que anteceden a la señal.

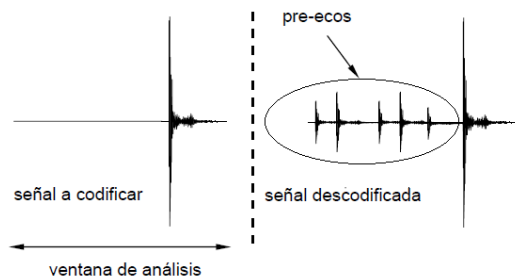


Figura 6-11. El efecto de los pre-ecos hace aparecer señal antes del impulso de la señal que los origina.

Este problema se agudiza en el caso de señales formadas por silencios seguidos de transitorios de nivel alto. Los codificadores por transformada son los que más sufren este tipo de problema, ya que necesitan grandes bloques de muestras para tener una gran resolución en frecuencia; los bloques pequeños limitan la resolución espectral.

Es esencia, los codificadores por transformada sacrifican la resolución temporal para tener una mejor resolución espectral. Los bloques grandes son apropiados para señales con variaciones lentas de nivel; sin embargo, cuando aparecen señales con transitorios altos es más conveniente utilizar bloques pequeños. Los codificadores por transformada presentan un retardo de procesamiento mayor que el de los codificadores sub-banda.

La mayoría de codificadores de baja tasa binaria emplean unos modelos psicoacústicos para cuantificar adaptativamente sólo las partes perceptualmente audibles de la señal. Las partes de la señal que se encuentran por debajo del umbral de audición (enmascaradas por las señales de nivel más alto) se consideran inaudibles y no se codifican. Además, la resolución de la cuantificación se adapta dinámicamente para hacer que el error de cuantificación esté cerca de las componentes más significativas del espectro, de manera que cuando se reconstruya la señal, el error sea enmascarado por ella. Esta técnica produce unas grandes reducciones de la tasa binaria. Los codificadores que trabajan en el dominio del tiempo pueden conseguir tasas de reducción de 4:1 y 12:1.

a) Codificación sub-banda

Las muestras de la señal de banda ancha se agrupan en pequeños bloques consecutivos y van pasando por un banco de filtros digitales, como se aprecia en la Figura 6-12. El banco de filtros divide la señal en múltiples canales paso-banda que tienden a asemejarse a las bandas críticas del oído. El banco de filtros debe tener frecuencias de corte con grandes pendientes de atenuación (en torno a los 10 dB/octava) con objeto de aproximarse a la respuesta que tienen realmente las bandas críticas y para confinar el ruido de cuantificación dentro de las mismas. Este tipo de diseño sólo se puede realizar con filtros digitales. Además, el retardo del filtrado (idealmente inferior a 3 o 4 ms) debe ser lo más pequeño posible para que el ruido de cuantificación no exceda la resolución temporal del oído. Las muestras de cada sub-banda son analizadas y comparadas entre sí según un modelo psicoacústico; el codificador cuantifica adaptativamente las muestras de cada sub-banda en base al umbral de enmascaramiento de la misma.

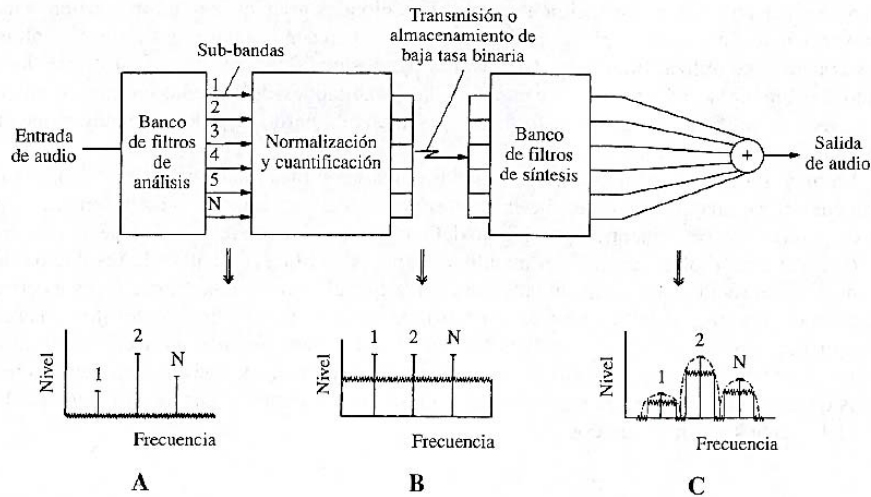


Figura 6-12. Los codificadores sub-banda crean bandas estrechas de muestras de alta resolución, con un bajo ruido de banda ancha. Las muestras en cada sub-banda se normalizan y se cuantifican sin mucha precisión, elevando el ruido de banda ancha. Cuando se reconstruyen las muestras de audio, los filtros modelan el ruido de cuantificación en cada banda, donde es enmascarado por la propia señal. **A:** Muestras sub-banda de alta resolución. **B:** Muestras sub-banda normalizadas y de reducida tasa binaria. **C:** Muestras de audio reconstruidas con el ruido de cuantificación enmascarado.

Cada sub-banda se codifica de forma independiente, con una mayor o menor asignación de bits. En cualquier caso, siempre se incrementa el ruido de cuantificación en cada una de las sub-bandas. Sin embargo, cuando se reconstruye la señal, el ruido de cuantificación dentro de cada una de ellas es enmascarado por las propias señales de audio de las sub-bandas, tal como se muestra en la Figura 6-12. Los niveles de ruido de cuantificación que antes podían ser molestos ahora son tolerados porque en las sub-bandas que haya señal, ésta se encargará de

enmascararlos. La asignación de bits se determina en base al modelo psicoacústico y al propio análisis de la señal; estas operaciones se realizan continuamente en cada sub-banda cada vez que entra un nuevo bloque de datos. Las muestras se cuantifican dinámicamente de acuerdo con la audibilidad de las señales y del propio ruido; en codificadores compatibles, existe una gran flexibilidad en el diseño del modelo psicoacústico y en la asignación dinámica de bits. Los decodificadores utilizan las muestras cuantificadas para reconstruir la señal de cada bloque; la señal de salida de banda ancha se reconstruye con un banco de filtros inversos, sumando las señales sub-banda.

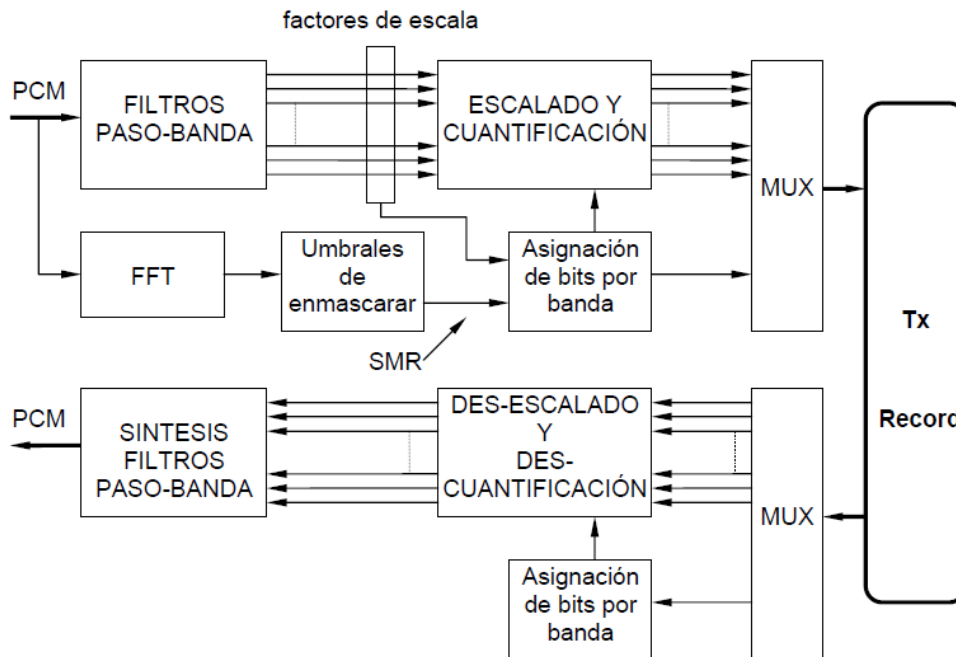


Figura 6-13. Esquema de funcionamiento de la codificación por sub-bandas. La señal pasa por M filtros paso-banda en paralelo, contiguos y con el mismo ancho de banda (f_s/M). La señal del filtro es diezmada por M y se re-muestrea al doble de su ancho de banda. Cada señal diezmada se cuantifica por separado. En el receptor a cada banda se le aumenta su frecuencia de muestreo por un factor M para finalmente sumarla todas.

Un codificador perceptual sub-bandas utiliza un banco de filtros digitales para dividir una señal de audio de corta duración en múltiples bandas, tal como se muestra en la Figura 6-14. En algunos diseños, la señal de audio pasa también por una transformada de Fourier (FFT) con objeto de analizar la energía en cada una de las bandas.

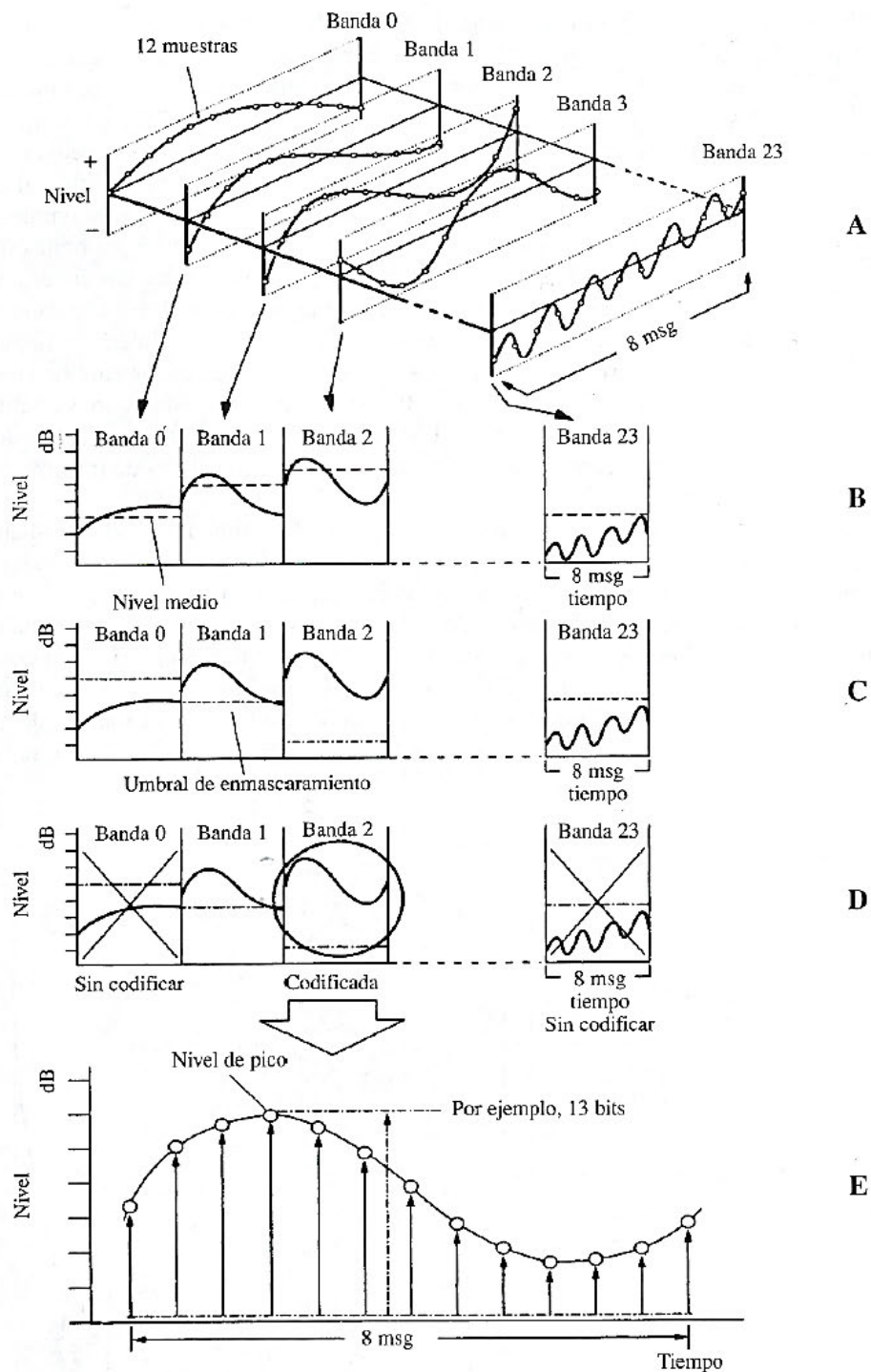


Figura 6-14. Un codificador sub-banda divide la señal en bandas estrechas de frecuencia, calcula el nivel medio de la señal y el umbral de enmascaramiento en cada una de ellas y después cuantifica de forma independiente las muestras de audio pertenecientes a cada sub-banda. A: Salida de un banco de 24 filtros. B: Cálculo del nivel medio en cada sub-banda.

C: Cálculo del umbral de enmascaramiento en cada sub-banda. D: Las sub-bandas por debajo del umbral de enmascaramiento no se codifican; las sub-bandas por encima del mismo sí se codifican. E: La asignación de bits en cada sub-banda se fija en función del nivel de pico por encima del umbral (Matsushita Electric⁵).

⁵ Actualmente el nombre de esta corporación es *Panasonic Electric*.

Estos valores son utilizados por el modelo psicoacústico para determinar la curva de enmascaramiento global que se ha de aplicar a la señal del bloque. Esto permite tener una mejor codificación en el dominio del tiempo. Específicamente, el codificador analiza la energía de cada sub-banda y determina cuáles de ellas contienen información audible. También se calcula la potencia media de cada sub-banda. Este valor se utiliza para determinar el nivel de enmascaramiento, debido al enmascaramiento que producen las señales de las bandas adyacentes. Finalmente, con los umbrales mínimos de audición de cada sub-banda se hallan definitivamente los umbrales de enmascaramiento. Los niveles máximos de señal en cada sub-banda se comparan con los umbrales de enmascaramiento. Las sub-bandas que no contengan información audible no se codifican. Análogamente, los tonos de una sub-banda que sean enmascarados por otros tonos cercanos tampoco se codifican, y en algunos casos sub-bandas enteras pueden llegar a enmascarar sub-bandas contiguas y, por tanto, éstas tampoco se llegarán a codificar. Las sub-bandas con niveles máximos de señal por encima del umbral de enmascaramiento son consideradas como bandas de información audible y deben ser codificadas. En la Figura 6-15 se muestran una serie de líneas verticales que representan los niveles máximos de potencia y los umbrales de audición y de enmascaramiento.

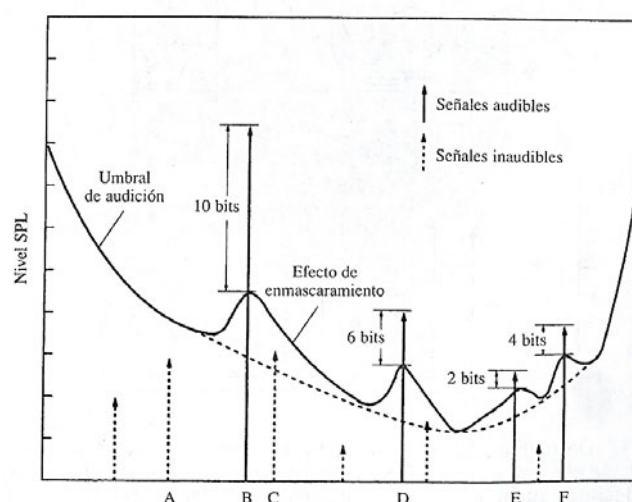


Figura 6-15. El algoritmo de asignación fija el número de bits de acuerdo con la audibilidad de las señales sub-banda. Los tonos inaudibles no tienen asignación de bits y por tanto no son codificados.

Las señales que estén por debajo del umbral de audición o el umbral de enmascaramiento no se codifican en ningún caso. Por ejemplo, la señal A está por debajo del umbral de audición y no sería codificada. La señal C tampoco se codificaría porque la señal B ha desplazado hacia arriba el umbral de enmascaramiento. En cambio las señales B, D, E y F sí que se codificarían.

b) Codificación por transformada

En la codificación por transformada, los bloques de muestras de audio se transforman al dominio de la frecuencia. El esquema de funcionamiento de este tipo de codificación se muestra en la Figura 6-16. Las transformadas que normalmente utilizan los codificadores son transformadas discretas de Fourier (DFT), como la transformada rápida de Fourier (FFT), o la transformada discreta del coseno modificada (MDCT). En estas transformaciones lo que se intenta aproximar es el análisis espectral de las vibraciones presentes en la membrana basilar a lo largo de su longitud. Los coeficientes espectrales obtenidos con las transformadas se cuantifican de acuerdo al modelo psicoacústico; las componentes enmascaradas son eliminadas, y la cuantificación se realiza siempre en base a la audibilidad. A diferencia de los codificadores sub-banda, que utilizan el análisis espectral para codificar muestras de la señal en el dominio del tiempo, los codificadores por transformada cuantifican los coeficientes espectrales de la señal. Grandes bloques de muestras proporcionan una gran resolución espectral pero se pierde en resolución temporal; por ejemplo, los bloques con una gran longitud pueden llegar a tener pre-ecos antes de un transitorio de nivel. La mayoría de codificadores trabajan con bloques solapados en tiempo al 50 % con objeto de mejorar la resolución temporal; la primera mitad de las muestras de un bloque corresponde a

la segunda mitad de las muestras del bloque anterior. Esta técnica permite disminuir las variaciones espectrales entre bloques.

Las muestras en el dominio del tiempo se transforman al dominio de la frecuencia, obteniéndose los coeficientes espectrales de la señal. A los números de los coeficientes espectrales se les denomina normalmente por el término 'bin'; por ejemplo, una transformada de 512 puntos produce 256 coeficientes o 256 bins. Los coeficientes, cuyo número puede ser de 512, 1024 e incluso más, son agrupados en 32 bandas emulando el análisis de bandas críticas que realiza el oído. Este espectro representa la misma información que tiene el bloque de muestras temporales. Los coeficientes espectrales de cada banda se cuantifican de acuerdo con el modelo psicoacústico del codificador; la cuantificación en cada banda puede ser uniforme, no uniforme, fija o adaptativa.

En los codificadores por transformada adaptativos se aplica un modelo para cuantificar adaptativa y uniformemente cada banda, pero todos los coeficientes dentro de una misma banda son cuantificados con un mismo número de bits. El algoritmo de asignación de bits se asegura de que la señal tenga una buena SNR para enmascarar al ruido. La asignación de bits suele seguir un proceso iterativo con objeto de distribuir todos los bits disponibles y aumentar el margen de codificación, manteniendo siempre una tasa binaria mínima. En algunos casos, la tasa binaria de salida por bloque puede ser fija o variable. Antes de proceder a la transmisión de los datos codificados, éstos son comprimidos por un proceso de codificación de entropía sin pérdidas como, por ejemplo, utilizando una codificación Huffman o una codificación de longitud variable. El decodificador cuantifica de forma inversa los coeficientes y aplica una transformada inversa para reconstruir la señal en el dominio del tiempo.

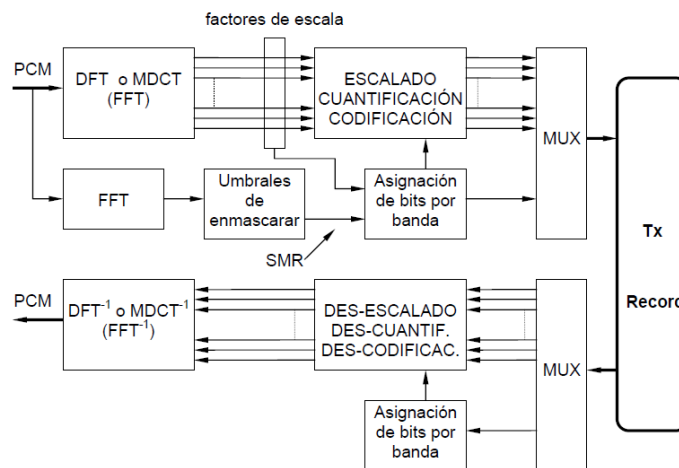


Figura 6-16. Esquema de funcionamiento de la codificación por transformada. La señal se divide en N puntos, para cada ventana de N puntos se realiza una transformada con N coeficientes, cuantificando cada coeficiente por separado. El decodificador realiza la transformada inversa de los coeficientes codificados.

Las transformadas DFT y DCT (MDCT) con cálculo FFT, que pueden considerarse bancos de filtros, son las encargadas de convertir la señal al dominio de la frecuencia. Con los coeficientes obtenidos se calcula la energía de la señal en cada una de las bandas críticas, lo que servirá para determinar el umbral de enmascaramiento en cada una de las bandas críticas. Los problemas que se pueden encontrar son chasquidos, provocados al unir distintas ventanas de datos, pre-ecos y baja calidad de la señal decodificada para regímenes binarios bajos.

Bancos de filtros

Los codificadores de baja tasa binaria suelen emplear un banco de filtros de análisis para dividir la señal de audio de banda ancha en sub-bandas más pequeñas; el decodificador utiliza un banco de filtros inversos para sintetizar las sub-bandas en una señal de banda ancha.

Un banco de filtros polifase es un conjunto de filtros de igual ancho de banda interrelacionados en fase y de eficiente implementación con buena resolución temporal y espectral, con atenuaciones cercanas a los 100 dB para tener un control perfecto del aliasing.

En un banco de filtros híbrido se emplea diferentes tipos de filtrado en cascada (como, por ejemplo, polifase y MDCT) para producir diferentes resoluciones espectrales a distintas frecuencias, con una complejidad de diseño asequible; por ejemplo, los codificadores MPEG layer III utilizan un banco de filtros híbrido que está formado por un banco de filtros polifase y por un conjunto de transformadas MDCT. En la Tabla 6-2 se comparan las propiedades de los distintos bancos de filtros empleados en los codificadores de baja tasa binaria.

Tabla 6-2. Comparativa de propiedades de bancos de filtros.

Características	Layer 1	Layer 2	Layer 3	AC-2	AC-3	ATRAC*	PAC/MPAC
Banco de filtros	PQMF	PQMF	Híbrido PQMF/ MDCT	MDCT/ MDST	MDCT	Híbrido QMF/ MDCT	MDCT
Resolución espectral a 48 kHz	750 Hz	750 Hz	41,66 Hz	93,75 Hz	93,75 Hz	46,87 Hz	23,44 Hz
Resolución temporal a 48 kHz	0,66 ms	0,66 ms	4 ms	1,3 ms	2,66 ms	1,3 ms	2,6 ms
Respuesta al impulso (LW)	512	512	1.664	512	512	1.024	2.048
Respuesta al impulso (SW)	—	—	896	128	256	128	256
Longitud de la trama a 48 kHz	8 ms	24 ms	24 ms	32 ms	32 ms	10,66 ms	23 ms

* ATRAC trabaja a una frecuencia de muestreo de 44,1 kHz. En esta comparativa, la longitud de trama y la respuesta al impulso del sistema ATRAC están dadas para una frecuencia de muestreo de 48 kHz (*Brandenburg y Bosi*).

6.5 Estándares de codificación

6.5.1 Estándar de audio MPEG-1

El grupo MPEG (*Moving Pictures Expert Group*) desarrolló en 1992 su primera norma de estandarización, la ISO/IEC 11172, “Codificación de imágenes en movimiento y audio asociado para la grabación digital hasta 1,5 Mb/s”, destinada a codificaciones de audio y vídeo digitales de tasa binaria baja. Este estándar es conocido con el nombre de MPEG-1. El estándar está formado por tres partes:

- El sistema (audio y vídeo multiplexado).
- El vídeo.
- El audio.

El máximo régimen binario para audio es de 1,856 Mb/s. La parte estándar dedicada al audio (ISO/IEC 11172-3) ha sido aplicada en el vídeo CD, el CD-ROM, la ISDN o RDSI, los videojuegos y la radiodifusión digital. Permite codificaciones de datos PCM con frecuencias de muestreo de 32, 44,1 y 48 kHz, a tasas binarias comprendidas entre 32 y 224 kb/s por canal (entre 64 y 448 kb/s para señales estéreo). Puesto que las redes digitales de datos utilizan tasas binarias de transmisión a 64 kb/s (8 bits a 8 kHz de frecuencia de muestreo), la mayoría de los codificadores presentan tasas binarias múltiplos de 64. El estándar ISO/MPEG-1 fue desarrollado específicamente para codificar audio y vídeo sobre formato CD (soporta 1,41 Mb/s). Además, en los modos estéreo se pueden explotar de forma opcional las redundancias e irrelevancias estereofónicas para reducir aún más la tasa binaria.

El sistema MUSICAM (*Masking-pattern Universal Subband Integrated Coding and Multiplexing*) es un algoritmo perceptual, basado en el sistema MASCAM (*Masking-pattern Adapted Subband Coding and Multiplexing*) sobre el que se desarrolló el MPEG-1. Este sistema divide la señal de audio en 32 sub-bandas y utiliza modelos de codificación perceptual, con umbrales de audición y enmascaramiento, para conseguir una

determinada reducción binaria. Con una frecuencia de muestreo de 48 kHz, cada una de las sub-bandas tiene un ancho de 75 Hz. Los resultados de pruebas del sistema MUSICAM muestran que con él se consigue una calidad equiparable a la del CD, que es totalmente compatible con señales monofónicas, que el uso de codificadores en cascada no produce degradación y que se consigue una calidad superior que en FM.

El estándar describe las tres capas de codificación, conocidas como *layers*:

- i. Layer I describe la codificación más simple y a unas tasas binarias relativamente altas (aproximadamente 192 kb/s por canal).
- ii. El layer II está basado en el layer I, aunque algo más complejo, y trabaja con unos regímenes binarios inferiores (entre 96 y 128 kb/s). El layer IIA es una versión *joint-stereo* (128 - 192 kb/s por cada pareja de señal estéreo).
- iii. El layer III es conceptualmente diferente de los layers I y II; es el más sofisticado y con las menores tasas binarias (aproximadamente 64 kb/s por canal). Es la capa que mejor calidad de audio presenta.

Con una tasa de 128 kb/s por canal se asegura la calidad de audio, luego se puede considerar que los layers II, IIA y III tienen una calidad aceptable de audio.

En términos muy generales, los tres codificadores trabajan de forma similar:

1. La señal de audio pasa por un banco de filtros y es analizada en el dominio de la frecuencia. Las componentes submuestreadas corresponden a los valores sub-banda o a los coeficientes espectrales.
2. La transformada de Fourier (o los datos de salida del banco de filtros) es utilizada para determinar los umbrales de enmascaramiento.
3. Los valores sub-banda o los coeficientes espectrales, según el caso, son cuantificados de acuerdo a un modelo psicoacústico.
4. Las muestras codificadas, conjuntamente con la información colateral de asignación binaria, se estructuran en tramas de datos antes de proceder a la transmisión.

Los codificadores no están definidos por el estándar MPEG-1, éste sólo especifica las características de los decodificadores. Esta asignación binaria adaptativa, abierta a futuras mejoras de codificación, especialmente del modelo psicoacústico, permite que los datos de salida de cualquier codificador perceptual sean decodificados según se especifica en el estándar.

Los layers del estándar MPEG-1 soportan codificación *joint-stereo* realizada con una codificación denominada codificación intensa. Las muestras de las sub-bandas de alta frecuencia de ambos canales (izquierdo y derecho) son sumadas entre sí formando un canal de audio conjunto; los factores de escala de cada uno de los canales se codifican de manera independiente. El decodificador reconstruye las envolventes de los canales izquierdo y derecho con sus respectivos factores de escala. La envolvente del espectro es idéntica en las sub-bandas superiores, tanto para el canal izquierdo como para el derecho, aunque las amplitudes no lo sean. La frecuencia a partir de la cual se codifica de forma conjunta en una codificación *joint-stereo* se selecciona entre cuatro posibles: 3, 6, 9 y 12 kHz, para una frecuencia de muestreo de 48 kHz pudiéndose modificar el valor de trama en trama.

Los datos se transmiten organizados en tramas, tal y como muestra la Figura 6-17, cada una de ellas con una entidad propia, necesaria para poder realizar correctamente la decodificación. La longitud en tramas depende del layer y del algoritmo MPEG. En MPEG-1, las tramas de los layers I y II tienen una misma longitud: 1152 muestras de audio. En el layer III los datos de audio de una trama están localizados en la trama de audio correspondiente. Además, en el layer III el número de bits por trama puede ser variable lo que permite cierta flexibilidad en función de las demandas de codificación que tenga la señal de audio en cada momento.

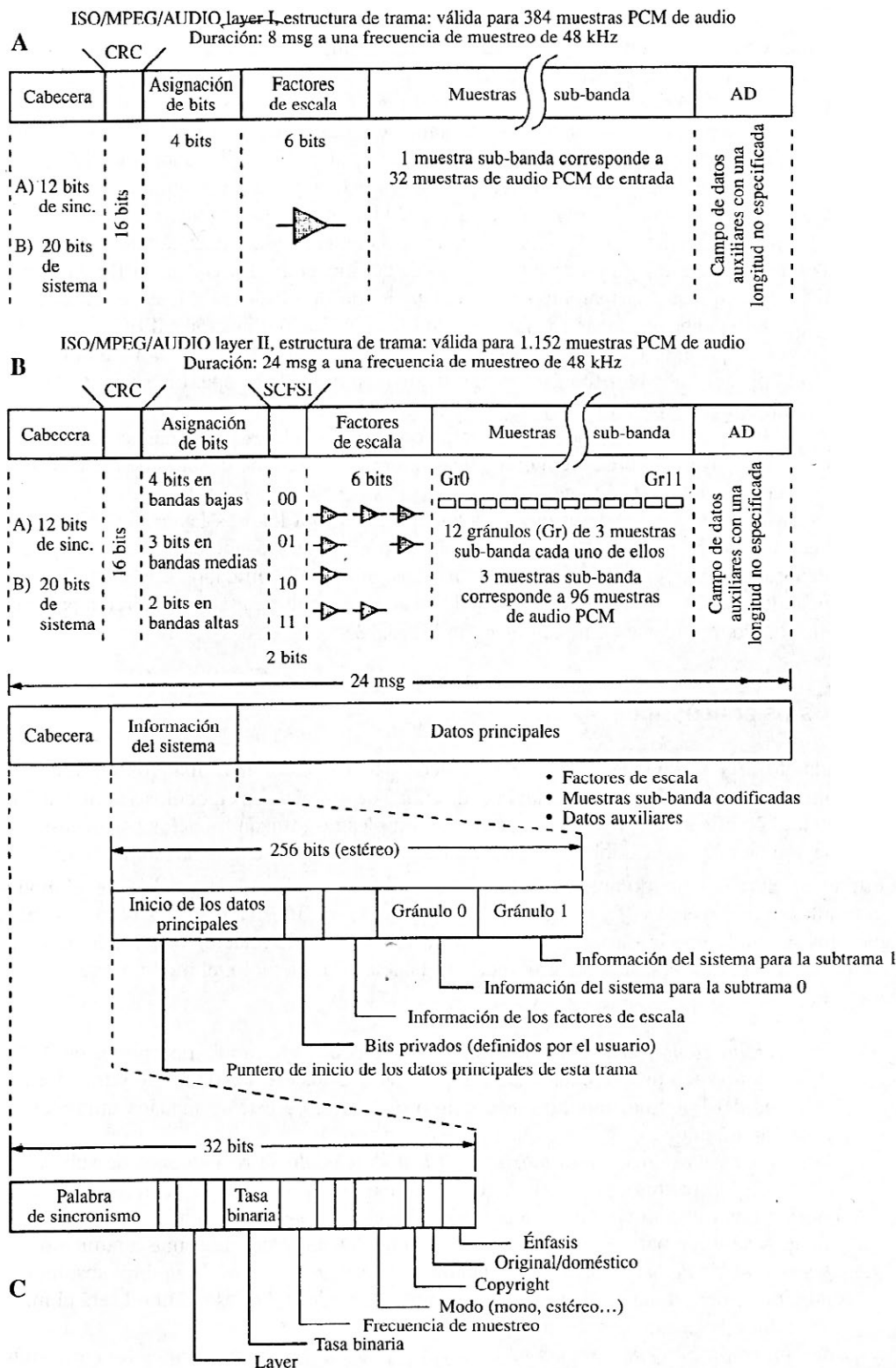


Figura 6-17. Estructura de tramas en los layers I, II y III del estándar ISO/MPEG-1. La cabecera y algunos otros campos son comunes, pero otros son diferentes. Los codificadores más avanzados pueden transcódecificar tramas de datos de un layer a otro. A: Formato de datos en el layer I. B: Formato de datos en el layer II. C: Formato de datos en el layer III.

La similitud que existe entre los layers permite que ambos puedan trabajar en tándem; por ejemplo, los datos del layer III se pueden convertir en datos codificados según el layer II sin pasar por el dominio analógico (aunque si es necesario realizar un determinado procesamiento digital). Un decodificador totalmente compatible con MPEG-1 puede decodificar los datos de su layer y de los layers inferiores. Los denominados 'codificadores X' sólo

pueden decodificar los datos de su propio layer. El layer I es el que mejor calidad tiene en aplicaciones de adquisición y producción de audio cuando se trabaja a altos regímenes binarios y la señal atraviesa seis o más procesos de codificación; con el layer II se puede realizar una correcta distribución de programas de audio cuando el número de procesos de codificación no es superior a dos; el layer III es el más eficiente y con los regímenes binarios más bajos, aunque con una calidad algo más baja; en este layer no es aconsejable realizar más de un proceso de codificación sobre la señal de audio.

Numerosas pruebas de evaluación han demostrado que tanto el layer II como el layer III a 2x128 kb/s, o a 192 kb/s en el modo *joint-stereo*, pueden ser utilizados para señales estéreo sin pérdida de calidad apreciable, comparada con la que ofrece un sistema lineal a 16 bits y a 44,1 kHz de frecuencia de muestreo. Si se utilizan tasas binarias más altas, como por ejemplo 384 kb/s, el layer I también presenta esta transparencia de calidad cuando se compara con una codificación lineal PCM de 16 bits. Con regímenes binarios de 128 kb/s, la calidad de las señales estéreo de los layers II y III se aproxima bastante a la transparencia (PCM 16 bits).

6.5.2 Estándar de audio MPEG-2

El estándar de audio MPEG-2, aprobado en noviembre de 1994, está diseñado para aplicaciones que van desde la televisión HDTV hasta la descarga de archivos vía internet. Este estándar multicanal es compatible con el MPEG-1 (estéreo), añade nuevas frecuencias de muestreo (frecuencias inferiores) y tiene un formato multicanal no compatible con MPEG-1, conocido como ACC.

MPEG-2 abarca todo el estándar MPEG-1, usando los mismos principios en los algoritmos de codificación y decodificación. Por ello ambos estándares son compatibles, un decodificador MPEG-2 acepta datos codificados con MPEG-1, un decodificador MPEG-1 puede reproducir una señal estéreo a partir de una señal multicanal codificada con MPEG-2; sin embargo, MPEG-2 también puede tener codificadores no compatibles con MPEG-1.

El estándar MPEG-2 codifica señales de audio multicanal a frecuencias de muestreo de 32, 44,1 y 48 kHz. Su diseño se ha realizado para que sea compatible con MPEG-1 (*backward compatible, BC*); normalmente se denomina MPEG-2 BC. La reducción binaria es bastante alta, debido a la redundancia que existe entre los 5.1 canales, de tal forma que estos pueden ser codificados de forma óptima con regímenes binarios comprendidos entre 384 y 640 kb/s. El estándar MPEG-2 también soporta codificaciones monofónicas y estéreo a frecuencias de muestreo de 16, 22,05 y 24 kHz, utilizando los layers I, II y III. En la Figura 6-18 se muestran todas las opciones de codificación de MPEG-1 y MPEG-2.

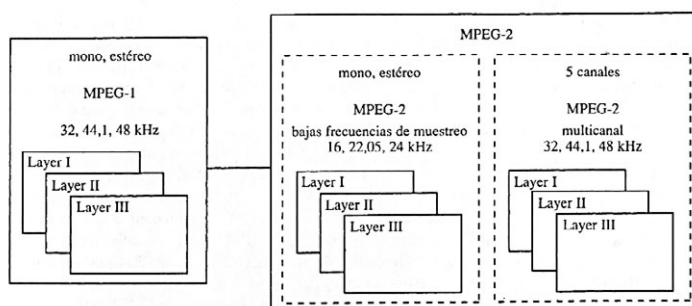


Figura 6-18. El estándar de audio MPEG-2 añade la posibilidad de codificar señales mono y estéreo con bajas frecuencias de muestreo, y una codificación multicanal. Admite los tres layers de MPEG-1.

El formato multicanal MPEG-2 utiliza básicamente cinco canales de audio, 3 frontales, 2 traseros envolventes (*surround*) y canal de graves (*subwoofer*). El formato multicanal MPEG-2 utiliza un codificador matricial para crear una mezcla estéreo, a partir de las cinco señales de audio, y que pueda ser decodificada por un decodificador estéreo MPEG-1. En la Figura 6-19 se muestra todo este proceso. Los canales izquierdo y derecho MPEG-1 son sustituidos por los nuevos canales izquierdo y derecho creados en la matriz del codificador MPEG-

2, con estructuras de tramas compatibles con MPEG-1. El resto de canales se colocan en la última parte de la trama de datos, en el campo auxiliar de datos. Un decodificador estéreo MPEG-1 ignora este campo auxiliar de datos y sólo reproduce los dos canales principales de audio. En algunos casos, en la matriz inversa de los decodificadores se pueden producir errores que se manifiestan como cancelaciones de señal en alguno de los canales, pero no así el ruido de cuantificación, que puede llegar a ser perfectamente audible.

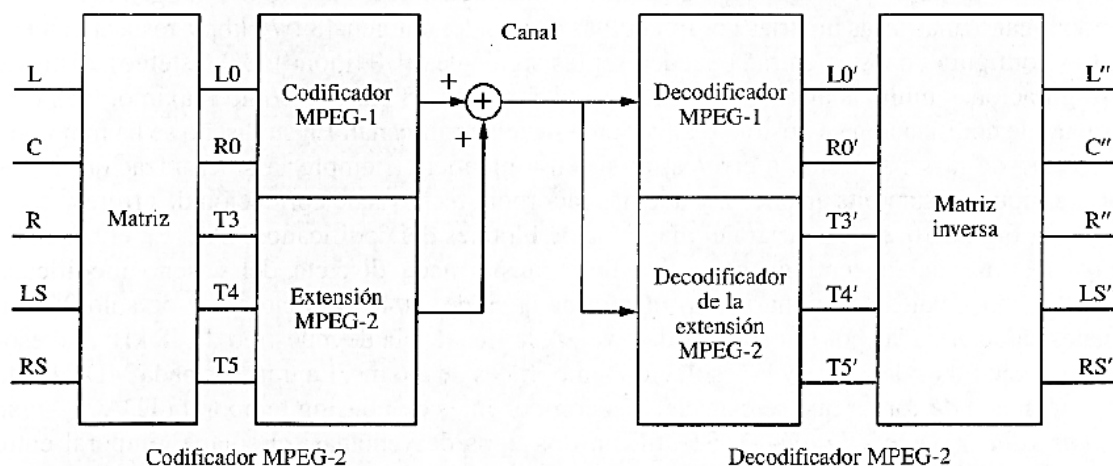


Figura 6-19. Codificador y decodificador de audio ISO/MPEG-2 en el que se muestra cómo se consigue un compatibilidad del formato multicanal 5.1 con el estándar de MPEG-1.

El estándar MPEG-2 utiliza las siguientes estrategias para obtener una tasa binaria de 384 kb/s (layer II y 48 kHz de frecuencia de muestreo): 'codificación intensa', atenuación de la diafonía intercanal, codificación predictiva intercanal y codificación de la imagen central fantasma. Las tasas binarias en MPEG-2 pueden llegar a un valor de 1066 kb/s. La trama de datos en MPEG-2 está dividida en dos partes. La primera parte es compatible con la trama MPEG-1, con una tasa máxima de 448 kb/s para el layer I, 384 kb/s para el layer II y 320 kb/s para el layer III. La segunda parte de la trama MPEG-2 contiene el resto de la información multicanal.

En el estándar MPEG-2 también se especifican los layer I, II y III a bajas frecuencias de muestreo (LSF) de 16, 22,05 y 24 kHz; con estas especificaciones se pierde la compatibilidad con los codificadores MPEG-1. Esta parte del estándar se la conoce por el nombre de MPEG-2 LSF. Con estos regímenes binarios tan bajos, el layer III es el que normalmente ofrece la máxima calidad de audio. Para adaptar este formato LSF al estándar MPEG-1 hay que realizar una serie de modificaciones en la tasa binaria y en las tablas de asignación de bits. La mejora en la calidad de audio es consecuencia de una mayor resolución espectral del banco de filtros polifase en bajas y medias frecuencias; esto permite tener un enmascaramiento más eficaz. En regímenes binarios muy bajos como 56 kb/s, las diferencias que presentan los layers I y II en señales de voz son mínimas. En la codificación LSF se utiliza el mismo flujo de datos y el mismo formato de trama. Para una frecuencia de muestreo de 24 kHz, la duración de la trama para el layer I es de 16 ms y de 48 ms para el layer II. En el layer III la duración de la trama se reduce en la misma proporción que lo hace el estándar MPEG-1.

El menú de tasas binarias, calidades y compatibilidades proporcionadas por MPEG se utilizan en un sinfín de aplicaciones, tales como en multimedia, CD-ROM, DVD-Vídeo, discos de ordenador, redes de área local, grabación y edición en estudios, grabación multicanal en disco duro, transmisión digital vía RDSI (ISDN), radiodifusión digital de audio y televisión digital multicanal. Existen numerosos programas implementados en C y C++ que realizan la codificación y la decodificación de los estándares MPEG-1 y MPEG-2, y que pueden ser descargados vía internet para ser ejecutados directamente en un ordenador personal (PC).

6.5.3 MPEG-2 AAC

El formato MPEG-2 AAC (codificación avanzada de audio) realiza una codificación estéreo o multicanal a una tasa binaria increíblemente baja de 64 kb/ por canal. También soporta una codificación multicanal 5.1 a una tasa binaria de 384 kb/s.

La codificación MPEG-2 AAC no es compatible con MPEG-1 (*non backward compatible* - NBC). La calidad obtenida por este formato de codificación es superior si se compara con el estándar MPEG-2 BC. MPEG-2 AAC tiene las típicas frecuencias de muestreo de 32, 44,1 y 48 kHz, y además permite otras nuevas dentro del margen 8 y 96 kHz, dando lugar a unas tasas binarias comprendidas dentro del margen 48 - 576 kb/s, respectivamente. Las configuraciones de entrada pueden ser las siguientes: mono, estéreo, distintas configuraciones multicanal hasta $3/2 + 1$ y provisión para 48 canales como máximo. Con este formato de codificación es posible realizar una mezcla multicanal. En su diseño se ha mejorado la calidad de la señal frente a errores; el sistema mantiene siempre la sincronización de los datos aunque existan bits erróneos, y además incorpora técnicas de corrección de errores.

En la Figura 6-20 se muestra el diagrama de bloques del codificador AAC. En el banco de filtros de la señal de entrada se utiliza una transformada discreta del coseno modificada (MDCT); en señales estacionarias la transformada es de 2048 coeficientes y de sólo 256 en señales caracterizadas por transitorios de nivel. A la frecuencia de muestreo de 48 kHz, la resolución espectral es de 23 Hz y la resolución temporal es de 2,6 ms. La transformada MDCT está implementada de forma que se cancelen las componentes de aliasing temporal (TDAC, *time domain aliasing cancellation*). Se utilizan dos tipos de ventanas; el solape temporal entre ventanas es del 50 %. Ambas ventanas, sinusoidal y Kaiser-Bessel, se utilizan en el banco de filtros; la ventana sinusoidal se utiliza cuando las componentes perceptualmente importantes tienen una separación espectral inferior a 140 Hz. La ventana Kaiser-Bessel se emplea cuando la separación entre componentes es superior a 220 Hz. La conmutación entre estos dos tipos de ventanas es de forma instantánea, incluso dentro del solapamiento temporal. Los coeficientes espectrales de la MDCT son filtrados con objeto de controlar la envolvente del ruido de cuantificación en cada ventana, realizando una ‘conformación temporal de ruido’ (TNS, *temporal noise shaping*). Este filtrado TNS se encarga de cambiar ciertas componentes espectrales por el residuo de predicción, el cual sirve para que el decodificador pueda controlar los pre-ecos, incluso dentro de la ventana del banco de filtros. El filtrado TNS permite tener una mejor codificación de las partes pseudoperiódicas de la señal, como en la señal de voz; las partes impulsivas de la señal de voz no siempre se codifican de forma óptima con una conmutación de ventanas convencional. Con el filtrado TNS, el codificador puede usar también bloques largos más a menudo sin introducir chasquidos esporádicos (*artifacts*).

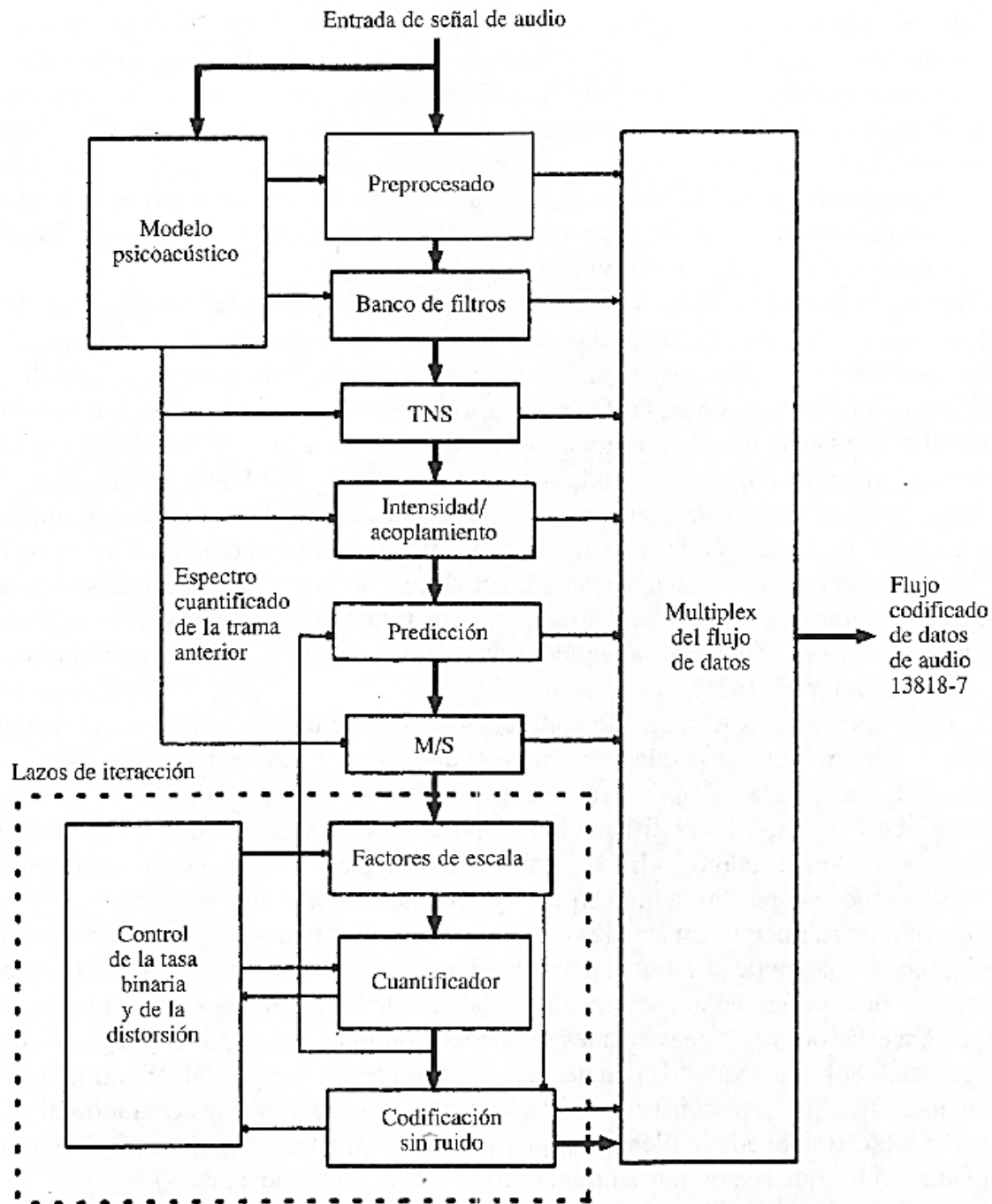


Figura 6-20. Diagrama de bloques del codificador MPEG-2 AAC. Las líneas más gruesas denotan el camino de los datos mientras que las líneas más delgadas denotan señales de control.

En pruebas de evaluación auditiva, el formato MPEG-2 AAC a 320 kb/s supera al estándar MPEG-2 layer II a 640 kb/s. Este último a 640 kb/s no es tan siquiera capaz de superar al AAC a 256 kb/s. MPEG-2 AAC asegura una calidad transparente para regímenes binarios comprendidos entre 256 y 320 kb/s, sin limitar el ancho de banda de los canales de audio. MPEG-2 AAC a 128 kb/s produce una mejor calidad que el MPEG-2 layer II a 192 kb/s o que el MPEG-2 layer III a 128 kb/s. MPEG-2 AAC a 96 kb/s es comparable al MPEG-2 layer II a 192 kb/s o al MPEG-2 layer III a 128 kb/s.

6.5.4 Codificador AC-3 (DOBLY DIGITAL)

El sistema de codificación AC-3 se utiliza ampliamente en señales de audio multicanal en aplicaciones de DVD-vídeo, DTV y DBS. Los sistemas de codificación anteriores al AC-3 fueron el AC-1 y AC-2.

El codificador estéreo AC-1 (*Audio Coding - 1*) es capaz de codificar una señal estéreo de 20 kHz a un régimen binario de 512 kb/s (aproximadamente una reducción de 3:1) y se usa en transmisiones vía satélite de televisión y en transmisiones FM, así como en servicios de radio vía cable.

El AC-2 es un sistema de codificación monocanal que se utiliza en aplicaciones de dos o más canales, todos ellos independientes entre sí. Utiliza una TDAC (*time domain aliasing cancellation*) con cancelación del aliasing temporal. El ancho de banda de la señal de audio se divide en sub-bandas con transformadas FFT de 512 puntos, solapadas al 50 %, alternando transformadas MDCT con MDST (transformadas discretas del seno y del coseno modificadas). La ventana que se aplica es del tipo Kaiser-Bessel. Los coeficientes espectrales se agrupan en sub-bandas (entre 1 y 15 coeficientes) para modelar los tamaños de las bandas críticas. El proceso de asignación de bits es adaptativo-regresivo; las asignaciones binarias se calculan tanto en el codificador como en el decodificador. Los coeficientes se codifican de acuerdo a la asignación de bits y después se agrupan en bloques. Los parámetros del algoritmo dependen de la frecuencia de muestreo. Para las frecuencias de muestreo de 48, 44,1 y 32 kHz, los parámetros de codificación son:

- a) Bytes/bloque: 168, 184, 190.
- b) Número total de bits: 1344, 1472, 1520.
- c) Sub-bandas: 40, 43, 42.
- d) Bits adaptativos: 225, 239, 183.

El codificador AC-2 tiene una alta calidad de audio a una tasa binaria de 256 kb/s por canal, aunque también puede trabajar a 128 y a 192 kb/s por canal. Los ficheros de audio codificados en AC-2 tienen una extensión específica registrada y son intercambiables entre ordenadores. Algunas aplicaciones del AC-2 son tarjetas de sonido de ordenadores, enlaces de transmisión entre estudios, enlaces digitales de larga distancia entre estudios vía RDSI (ISDN) en sesiones de grabación. El flujo de datos AC-2 es bastante robusto frente a errores. Dependiendo de la aplicación, el retardo en AC-2 varía de 7 a 60 ms.

El sistema de codificación AC-3 (conocido como Dolby Digital) es la culminación del formato AC-2, destinado concretamente a salas comerciales de cine. Es un codificador perceptual diseñado para procesar una serie de canales de audio; puede codificar de 1 a 6 canales de audio con diferentes configuraciones incluyendo un canal de baja frecuencia. Normalmente, AC-3 proporciona una configuración multicanal envolvente 5.1: canal izquierdo, derecho, central, izquierdo envolvente, derecho envolvente y el canal de efectos en bajas frecuencias. Los seis canales de audio ($6 \times 48 \text{ kHz} \times 18 \text{ bits} = 5184 \text{ Mb/s}$ en una representación PCM) se pueden codificar a una tasa nominal de 384 kb/s, lo cual supone una reducción binaria del orden de 13:1. El estándar AC-3 sólo soporta regímenes binarios comprendidos entre 32 y 640 kb/s. Es compatible con otros formatos previos de sonido multicanal, estéreo y monocanal; todos estos formatos pueden ser extraídos directamente de la estructura de datos AC-3. El estándar AC-3 no utiliza una estructura matricial 5.1 en su flujo de datos, lo cual es muy significativo; esto impide que el ruido de cuantificación pueda dirigirse a un canal equivocado, donde no pueda ser enmascarado. El AC-3 transmite un flujo de datos de audio multicanal, con información de la mezcla multicanal que se debe realizar en el decodificador (mono, estéreo, matriz envolvente o completamente multicanal).

Además, el AC-3 incorpora un control de normalización de nivel de los diálogos de manera que éstos, al ser reproducidos, mantengan un nivel constante en todos los programas y en todos los canales de audio. Con esta normalización, el usuario puede seleccionar el nivel de presión sonora y el decodificador automáticamente reproducirá todo el diálogo a ese mismo nivel, independientemente de cómo fue realmente grabado. El sistema también incluye un control del margen dinámico de la señal, cuya información va incluida en la trama de datos de manera que el decodificador pueda variar el margen dinámico de la señal codificada en $\pm 24 \text{ dB}$. También se

pueden incluir servicios auxiliares como descripciones verbales para invidentes, diálogos con una realzada inteligibilidad para las personas con deficiencias auditivas, comentarios y un segundo programa estéreo.

El AC-3 utiliza un algoritmo de asignación binaria adaptativo, regresivo/progresivo, que funciona tanto en el codificador como en el decodificador. El modelo psicoacústico se encarga de determinar la envolvente espectral, la cual va codificada en la estructura de datos. En la Figura 6-21 se muestran los diagramas de bloques del codificador y del decodificador AC-3.

La codificación AC-3 consigue una reducción binaria superior a la AC-2, ya que elimina todas las redundancias existentes entre los distintos canales de audio; en la versión estéreo de AC-3 se obtiene una calidad excelente a 192 kb/s. El formato AC-3 también suministra: datos relativos al programa de audio (mono, estéreo, matricial, etc.), datos de compresión dinámica, datos de encaminamiento de la señal a efectos de los *subwoofers* y datos de control de ganancia del programa de audio. El sistema AC-3 puede trabajar con muestras de 24 bits en un ancho de banda de 20 Hz – 20 kHz; el ancho de banda del canal de efectos se extiende únicamente entre 20 - 120 Hz. Las frecuencias de muestreo pueden ser de 32, 44,2 y 48 kHz. Los regímenes binarios del sistema AC-3 están comprendidos entre 32 kb/s (señal mono) y 640 kb/s (señal multicanal 5.1), con un gran abanico de aplicaciones. Con AC-3 es posible realizar una edición de audio a nivel de bloque, pudiendo reproducir la señal de audio en el decodificador tanto en sentido normal como en sentido inverso (lectura de bloques de audio en uno y en otro sentido).

El sistema de codificación AC-3 es popularmente conocido por el nombre de Dolby Digital y se utiliza en una gran cantidad de aplicaciones. Dolby Digital proporciona 5.1 canales de audio de codificación multicanal para salas de cine (el formato AC-3 se introdujo por primera vez en la película “El regreso de Batman”). Los datos están impresos de forma óptica entre los dientes de la cinta, a una tasa binaria aproximada de 320 kb/s. Las pistas analógicas quedan intactas; el sistema AC-3 es totalmente compatible con los sistemas analógicos de audio. El sistema Dolby Digital Surround EX añade un canal central envolvente (*surround*); este nuevo canal es mezclado matricialmente con los canales envolventes izquierdo y derecho. Se puede obtener fácilmente en el decodificador añadiendo un pequeño circuito.

En aplicaciones de consumo, Dolby Digital se utiliza para codificar audio multicanal 5.1 (o un número inferior) en sistemas de transmisión por cable y vía satélite y en productos domésticos de vídeo como el DVD-vídeo. El sistema Dolby Digital es un formato opcional del estándar DVD-Audio. También se utiliza para codificar el audio de la televisión digital estándar ATSC DTV; fue seleccionado en 1993 por el comité ATSC.

Mientras que en una codificación tradicional de audio la calidad del sistema está relacionada en términos de especificaciones y medidas, en una codificación perceptual la calidad se define en términos de percepción y fisiología. La tecnología tradicional de audio la podemos encontrar en cualquier disciplina de la técnica, pero los codificadores perceptuales sólo se encuentran en el ámbito del mundo de la psicoacústica.

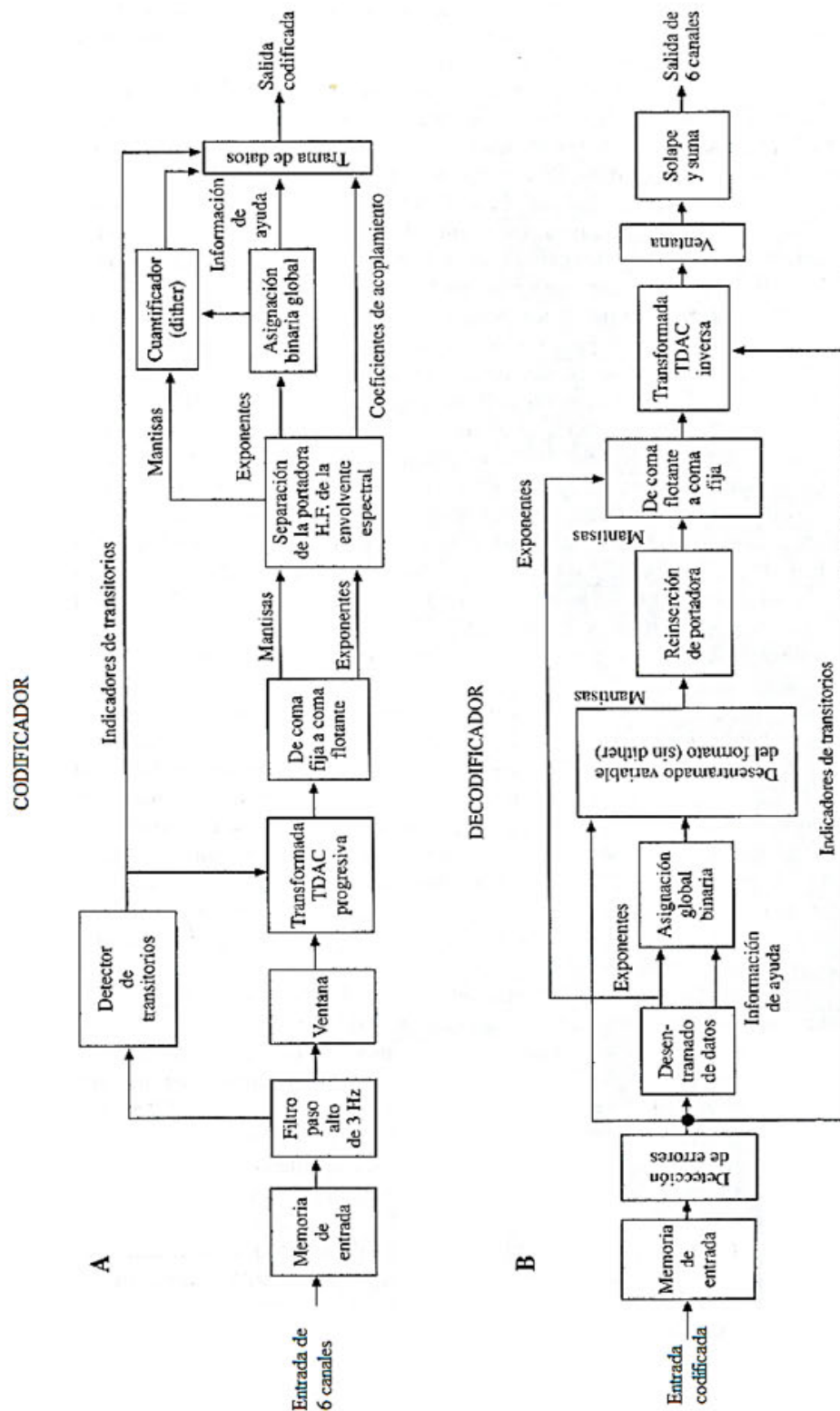


Figura 6-21. Codificador y decodificador de transformada adaptativa AC-3. Este codificador proporciona un sonido multicanal 5.1. A: Codificador. B: Decodificador (Laboratorios Dolby).

6.6 Cuestionario

1-En ausencia de sonido alguno, nuestro umbral de enmascaramiento hace que la mayor sensibilidad de percepción auditiva la tengamos en:

- a) Bajas frecuencias (<100 Hz)
- b) Frecuencia medias (1-2 kHz)
- c) Altas frecuencias (>10 kHz)
- d) En bajas y altas frecuencias

2- A grandes rasgos se podría definir a nuestro sistema auditivo como un conjunto de filtros paso banda de un tercio de octava.

- a) Verdadero.
- b) Falso.

3- El enmascaramiento frecuencial:

- a) Ocurre cuando un tono desplaza la curva umbral en un margen de frecuencias circundante (frecuencias más altas).
- b) Ocurre cuando un tono desplaza la curva umbral en un margen de frecuencias circundante (solo en frecuencias más bajas).
- c) Es una deficiencia auditiva.
- d) Ninguna de las anteriores es cierta.

4- El ruido de cuantificación será audible siempre que dicho nivel sea inferior al umbral de enmascaramiento: $SNR < SMR$.

- a) Verdadero.
- b) Falso.

5- El enmascaramiento temporal aparece cuando los tonos se suceden rápidamente, pero no de forma simultánea y el oído los integra como uno solo.

- a) Verdadero.
- b) Falso.

6- Los codificadores perceptuales eliminan la información irrelevante y redundante de la señal de audio sin que nuestro sistema auditivo sea capaz de percibirlo.

- a) Verdadero.
- b) Falso.

7-La diferencia entre la codificación por sub-bandas y la codificación adaptativa por transformada es:

- a) El filtro FIR predictor.
- b) Por sub-banda se procesa en el dominio del tiempo, mientras que por transformada se hacen en el dominio de la frecuencia.
- c) Por transformada se procesa en el dominio del tiempo, mientras que por sub-banda se hace en el dominio de la frecuencia.
- d) Ninguna de las anteriores es cierta.

8-Los bancos de filtros en cuadratura:

- a) Sirven únicamente en la codificación por transformada.
- b) Siempre reconstruyen la señal sin *aliasing* a partir de N sub-bandas solapadas en frecuencia.
- c) Sirven para la codificación por sub-bandas.
- d) Ninguna de las anteriores es cierta.

9-La Layer-I admite una tasa binaria de:

- a) Aproximadamente 192 kbps/canal.
- b) 64 kbps/canal.
- c) Menos de 128 kbps/canal.
- d) Ninguna de las anteriores es cierta.

10-La codificación *joint-stereo*:

- a) Las muestras de las sub-bandas de alta frecuencia de ambos canales (izquierdo y derecho) son separadas formando un canal de doble de audio conjunto
- b) Es soportada únicamente por el estándar MPEG-2 debido a su complejidad en la codificación.
- c) Las muestras de las sub-bandas de alta frecuencia de ambos canales (izquierdo y derecho) son sumadas entre sí formando un canal de audio conjunto
- d) Ninguna de las anteriores.

11-Las tramas del MPEG-1:

- a) Contienen 384 muestras en el layer I y 1152 muestras en el layer II y III.
- b) Contienen 1384 muestras en el layer I y 1152 muestras en el layer II y 384 muestras en el layer III.
- c) Terminan con una cabecera ISO de 32 bits.
- d) Ninguna de las anteriores es cierta.

12- ¿Dónde se realiza una transformación espacio-frecuencia del sonido?

- a) Membrana basilar
- b) Estribo
- c) Oído Medio
- d) Ninguna de las respuestas anteriores

13- Un golpe de batería de gran nivel puede enmascarar

- a) A señales débiles que le precedan
- b) A señales periódicas que le precedan
- c) A señales aleatorias que le precedan
- d) A señales de voz que le precedan

14-En un codificador DPCM:

- a) Existe un filtro FIR adaptativo que actúa como predictor
- b) Se mejora 12 dB el parámetro SNR

- c) Se consigue una reducción de 4:1
- d) Ninguna de las respuestas anteriores.

15- En un codificador MPEG-1, layer II, la duración de la trama para una frecuencia de muestreo de 32kHz es:

- a) 8ms
- b) 24ms
- c) 48ms
- d) Ninguna de las anteriores

16- En un codificador perceptual por transformada la división espectral de la señal se realiza con:

- a) 32 filtros paso banda.
- b) 1 transformada modificada del coseno.
- c) 16 transformadas modificadas del coseno.
- d) 32 transformadas modificadas del coseno.

17- La codificación “transparente” para MPEG-1 Layer III se obtiene para una reducción binaria:

- a) 4/1. → Para MPEG-1 Layer I
- b) 8/1. → Para MPEG-1 Layer II
- c) 12/1.
- d) 10/1.

18- El CRCC en una trama de datos nos permite:

- a) Detectar y corregir errores.
- b) Sólo corregir errores.
- c) Sólo detectar errores.
- d) Ninguna de las anteriores.

19- ACC realiza una división espectral mediante

- a) Transformadas de coseno (MDCT).
- b) Transformadas de seno (MDST).
- c) Transformadas alternada de coseno y seno (MDCT-MDST).
- d) Transformadas de Fourier.

20- La codificación MPEG 2 AAC es un tipo de codificación:

- a) Sub-bandas.
- b) Transformada.
- c) Híbrida.
- d) Ninguna de las anteriores.

7 CODIFICACIÓN DE LA SEÑAL DE VÍDEO

Las tecnologías digitales implican importantes ventajas en el tratamiento de las señales que suministran información de imagen, pero no ha sido posible su utilización por el gran público hasta la aparición de potentes métodos de compresión. En el campo profesional y bajo la denominación de Rec. 601 de UIT (Unión Internacional de Telecomunicación) se encuentra desde hace tiempo una Norma que regula la televisión digital dentro de los estudios de televisión. Esta digitalización ya se ha visto en el apartado de digitalización de vídeo. Este tipo de digitalización entraría dentro de la codificación de fuente, donde no hay compresión de la señal generando regímenes binarios muy altos.

En la actualidad y gracias a los algoritmos de compresión MPEG, ya es posible la transmisión digital de señales de vídeo, puesto que se ha reducido drásticamente la velocidad binaria necesaria y por lo tanto el ancho de banda que necesita el canal de transmisión.

Las ventajas de la televisión digital no se encuentran concentradas únicamente en la mejora de la calidad de la imagen, sino en la gran cantidad de servicios añadidos que puede ofrecer. Una de las importantes posibilidades es la interactividad entre el usuario y el radiodifusor o broadcaster, como por ejemplo en transmisiones deportivas, donde se puede ofrecer la posibilidad de que desde casa se elija cuál de las cámaras se quiere visualizar, de entre las existentes en el estadio. También se permite disponer de un enorme catálogo de películas y material audiovisual para poder ser visualizado, incluso bajo demanda.

Alrededor de este tema se han formado diferentes grupos de estudio y trabajo encargados de desarrollar los algoritmos de compresión (MPEG *Moving Picture Experts Group*) y las técnicas de emisión de la señal de vídeo (DVB *Digital Video Broadcasting*).

En el año 1988 se creó un Comité de MPEG con el fin de desarrollar una técnica que fuese capaz de codificar digitalmente señales de vídeo con su audio asociado, sin que la tasa binaria resultante superase valores del orden 1,5 Mb/s. Este grupo desarrolló una técnica que dio origen a un estándar de nombre ISO/IEC 11172, conocida popularmente como MPEG-1. Más tarde y como mejora de este se desarrolló otra norma, la ISO/IEC 13818, conocida como MPEG-2.

La norma MPEG-2 está dividida en 3 bloques:

1. ISO/IEC 13818-1: parte de la norma que describe la denominada Trama de Sistema, donde se analiza la forma de multiplexar los muchos canales posibles (de vídeo, de audio y de datos), para conformar el Flujo de Transporte, **TS (TRANSPORT STREAM)**, así como la información de temporización que permite al decodificador efectuar la correspondiente sincronización.
2. ISO/IEC 13818-2: en la que se analiza con todo detalle la denominada **Trama de Vídeo, ES Video (Elementary Stream)**, y es donde se establece la técnica para comprimir un único canal de vídeo (sin su audio asociado) y formar con él un flujo binario que lo identifica.
3. ISO/IEC 13818-3: denominada **Trama de Audio, ES Audio (Elementary Stream)**, y como su nombre indica, describe los procesos para comprimir y codificar los canales de audio que van acompañando a la señal de vídeo.

Diferencias entre MPEG-1 y MPEG-2

La norma MPEG-2 se desarrolló a partir de la MPEG-1, por lo que es muy compatible con ella. Esto quiere decir que un decodificador de un tipo, puede entender relativamente bien la señales codificadas con el otro, si bien la trama MPEG-2 presenta ciertas mejoras frente a su antecesora, lo que origina que los decodificadores MPEG-1 no podrán entender algunas variables recibidas y por tanto no procesarán esa señal MPEG-2. En cuanto a la trama de vídeo, pueden resaltarse dos importantes diferencias:

- En MPEG-2, si se quiere, puede utilizarse el formato de exploración entrelazada, lo que no es posible utilizar en MPEG-1, donde solo existe exploración progresiva. Por compatibilidad y universalidad, MPEG-2 también soporta el formato de exploración progresiva.
- MPEG-2 puede soportar diferentes niveles de resolución en la señal transportada, lo que no es posible en MPEG-1. Esto se ha hecho pensando en la televisión comercial presente y futura, para que la norma no se quede limitada a una única resolución, que puede resultar obsoleta en los próximos años.

Los tamaños de las imágenes pueden adoptar cualquier valor, siendo suficiente que al principio de la secuencia se indique cual es el valor de dicho tamaño. En la Tabla 7-1 se muestran los tamaños que se aceptan como utilizables en la actualidad.

Tabla 7-1. Parámetros de los distintos estándares de televisión.

Formato	Imag/s	Puntos x línea	Mb/s
SIF	12,5	352 x 288	1,2 – 3
SDTV	25	720 x 576	4 – 10
EDTV	25	960 x 576	6 – 15
HDTV	25	1920 x 1152	20 – 40

Este aspecto no debe confundirse con la calidad de la imagen, puesto que para cada uno de los tamaños anteriores puede regularse adicionalmente el nivel de calidad. Esta posibilidad de regular la calidad de la imagen codificada se puede llevar a cabo de dos formas diferentes:

- Modificando la velocidad binaria del flujo que la transporta. Cuanto mayor sea el régimen binario utilizado para el transporte mayor será la calidad obtenida en la imagen transportada.
- Mediante la técnica denominada escalabilidad, consistente en codificar varias capas, cada una de ellas aportando informaciones adicionales sobre las anteriores. En los decodificadores es posible sacrificar calidad, tomando solamente una parte de la trama y olvidándose del resto.

En MPEG-2 se sigue manteniendo la característica de que el proceso de codificación-decodificación es asimétrico. Esto quiere decir que se complica todo lo posible el codificador con la idea de que se simplifique al máximo el decodificador, lo cual es lógico si se piensa que codificador solo hay uno y decodificadores existen o pueden existir muchos y hay que abaratar su coste.

La norma define las características que debe poseer el flujo binario que lleva las informaciones y por lo tanto especifica las características de los decodificadores.

7.1 Principios de compresión

El número de datos generados por el conversor analógico-digital de las señales de vídeo es tan elevado que su tratamiento directo para transmisión sería impracticable, y su codificación resultaría tan compleja y extensa que no se podría llevar a cabo. Por ello es tan importante la compresión en el proceso de digitalización de la señal de vídeo y de audio. Ésta permite a los sistemas digitales usar un menor ancho de banda o lo que es equivalente un menor régimen binario, y por lo tanto, una menor capacidad de almacenamiento con mejores características y a menor precio que los dispositivos analógicos. De ahí la importancia de la compresión ya que sin ésta la digitalización no supondría una mejora respecto a los sistemas analógicos.

7.1.1 Compresión con pérdida y sin pérdida

El mayor avance de un sistema digital es que es posible la perfecta recuperación de datos (y la información que estos representan). Esto se debe a la compresión sin pérdidas, lo cual significa que los métodos de compresión no causan pérdida alguna de datos ni generan ningún tipo de error. El suprimir datos y después poder recuperarlos sin error alguno es posible gracias a que prácticamente todos los datos reales contienen patrones repetidos que el compresor detecta, ordena y transmite de manera más eficiente. En general los algoritmos usados en la compresión sin pérdidas no dependen de la información que se está tratando, puede trabajar con cualquier tipo de dato, pero el grado de compresión conseguido sí que puede depender de ésta.

Dentro de los métodos de compresión sin pérdida podemos encontrar la codificación *Run-Length*, la codificación estadística o la compresión LZW (Lempel-Ziv-Welch).

A pesar de las ventajas de la compresión sin pérdidas la mayoría de sistemas digitales de audio y vídeo usan sistemas de compresión con pérdidas. Esto significa que parte de los datos son modificados o directamente eliminados. Sin embargo y a pesar del nombre, la pérdida de información no supone siempre una pérdida de calidad para el observador, ya que los datos que han sido modificados o eliminados pueden no afectar de manera significativa a la señal, con lo que ésta, después de la descompresión, apenas se verá afectada. El grado de pérdida de datos puede modificarse lo cual puede dar como resultado una compresión altísima y una señal de salida de baja calidad o una compresión considerable y una señal de salida prácticamente igual a la de entrada al sistema.

De los métodos con pérdidas se pueden destacar el truncamiento, el sub-muestreo, la codificación diferencial, la codificación predictiva, la codificación por transformada o la compensación de movimiento.

7.1.2 Redundancia

En referencia a la compresión con pérdidas, la redundancia de los datos de audio y de vídeo son tales que ayudan a reducir los bits usados para transmitir la información, sin que esto produzca pérdida perceptible al observador.

a) Redundancia espacial

Es la correlación estadística existente entre píxeles que están próximos entre sí. El valor de un píxel está fuertemente relacionado con los que tiene en sus proximidades dentro de la imagen. Gracias a esto no es necesario tratar cada píxel independientemente de los demás, sino que se pueden usar bloques enteros de píxeles, así en MPEG y JPEG (codificación para imágenes fijas) por ejemplo, la imagen es dividida en bloques de 8x8 píxeles a los que se les aplica la DCT (Transformada Discreta del Coseno), lo que reduce de manera significativa el número de operaciones que ha de realizar el sistema.

b) Redundancia temporal

Es la correlación estadística entre píxeles cercanos en tiempo. Esto se explica de la siguiente manera: un píxel de un cuadro al siguiente puede haber variado muy poco o nada luego sólo será necesario saber con precisión cuál ha sido esa variación, en caso de que se hubiese producido. En MPEG por ejemplo, en lugar de codificar las imágenes en sí, se codifica solo el movimiento de ésta, así como las diferencias entre una imagen y la que le precede. También se usa la predicción y por tanto, la interpolación entre valores pasados y futuros.

c) Redundancia de codificación

Hay símbolos de la fuente que van a ser repetidos más veces que otros, luego a éstos se les puede asignar códigos más cortos que a los símbolos con menos probabilidad de salir. Así se consigue reducir el régimen binario de la codificación. Estos símbolos pueden ser píxeles, bloques de píxeles, coeficientes de transformadas, etc. El código *Huffman* o el código *Run-Length* son ejemplos del aprovechamiento de este tipo de redundancia.

Huffman

Se basa en un análisis estadístico de los códigos a transmitir, asigna códigos cortos a los valores más frecuentes y códigos largos a los menos probables. El rendimiento varía con el algoritmo particular y el tipo de imagen, pero pocas veces pasa la compresión de 8:1. Tiene un peor comportamiento para ficheros que contienen cadenas largas de píxeles idénticos, que suelen comprimirse mejor con el *Run-Lenght* u otro tipo.

Se necesitan estadísticas precisas de repetición de valores, codificación en dos pasadas, primera para el modelo estadístico y en la segunda los códigos de compresión variable lo que lo convierte en un método lento y muy sensible a la pérdida o adición de bits.

Run-Length

Funciona correctamente en imágenes con largas cadenas de valores repetidos (áreas de grandes sombras uniformes y matices constantes, imágenes artificiales). Es un sistema simple de construir e implementar que se basa en sustituir la cadena por valores simples y un contador.

d) Redundancia Psicovisual

La iluminación de una región, tal como es percibida por el ojo, depende de otros factores además de la luz reflejada por esta región. Este fenómeno se produce porque el ojo humano no responde con la misma sensibilidad a toda la información visual. Cierta información simplemente tiene menor importancia relativa que otra en el proceso visual normal. Se dice que esta información es *psicovisualmente redundante*, y se puede eliminar sin que se altere significativamente la calidad de la percepción de la imagen. Dado que hay información no relevante para el ojo, ésta no ha de ser tratada como aquella en la que éste si es muy sensible. La cuantificación no uniforme es un ejemplo de aprovechamiento de la redundancia psicovisual: escalones pequeños para la información que apreciamos al detalle (frecuencias bajas) y escalones mayores para la información a la que nuestro ojo es menos sensible (frecuencias altas).

7.2 Compresión y codificación MPEG

7.2.1 Descomposición de la imagen

Cada una de las imágenes que forman parte de la secuencia de vídeo, se encuentra configurada por píxeles, siendo su número función de la calidad que se pretenda conseguir en la imagen. Cuanta más calidad se quiera conseguir, mayor será el número de píxeles a usar.

En MPEG la imagen se descompone en un conjunto de pequeñas subimágenes denominadas macrobloques y el tamaño de estos es de 16x16 píxeles.

Se ha elegido este tamaño de 16x16 en base a varios criterios, entre los cuales pueden considerarse:

El tamaño debe ser una potencia de 2, teniendo en cuenta que el procesado será digital y binario y cumpliendo esta característica se aprovechan mejor los recursos.

Los macrobloques se van a descomponer posteriormente en bloques y el tamaño de estos también debe ser potencia de 2. Dado que la crominancia se va a submuestrear por 2, el tamaño de 8x8 es un valor muy apropiado.

Interesa que el grado de parecido entre todos los píxeles del macrobloque sea muy alto, pues como se verá posteriormente, cuanto más parecidos sean entre sí, más se podrá comprimir el macrobloque. En cuanto a este aspecto podría configurarse el macrobloque en 8x8 píxeles, para obtener bloques de 4x4, pero esto daría lugar a un número de macrobloques enorme para cada imagen y como cada uno de ellos lleva o puede llevar una cabecera, se ha llegado a la conclusión que sopesando ventajas e inconvenientes, el valor de 16x16 es el óptimo para los macrobloques.

Desde el punto de vista didáctico, el mismo proceso que se sigue para un macrobloque es el mismo que se sigue para otro, por lo que a partir de este momento toda la atención debe estar concentrada en la técnica utilizada para comprimir un macrobloque. Por tanto, cada imagen, sea del tamaño que sea, será descompuesta en macrobloques, tal como puede apreciarse en la Figura 7-1.

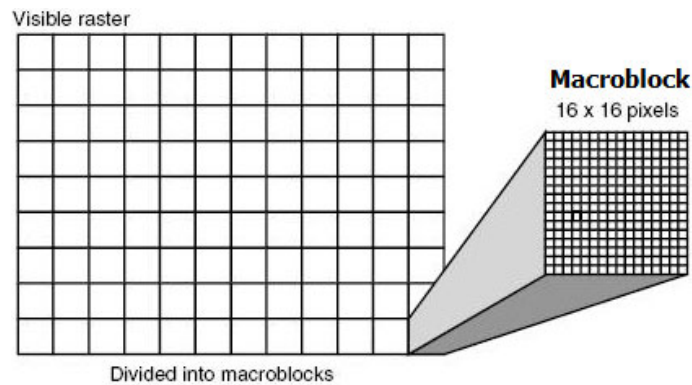


Figura 7-1. Descomposición de la imagen en macrobloques.

7.2.2 Concepto macrobloque

Cada macrobloque es un conjunto de 16x16 píxeles, pero de cada uno de ellos se van a obtener tres señales, por lo que a su vez, como se verá posteriormente, cada macrobloque se va a descomponer en un número determinado de bloques. El número de bloques por macrobloque no es fijo, sino que depende de algunos aspectos que serán considerados en los siguientes apartados.

Un bloque es una matriz de 8x8 números (64 valores) y es lo que configura la unidad básica de procesado. En la Figura 7-2 se puede apreciar la descomposición de un macrobloque en bloques.

Llegados a este punto, puede decirse que el mismo proceso seguido por un bloque servirá para el resto de los bloques, por lo que toda la atención deberá estar concentrada en la actuación para un bloque, porque el proceso será repetitivo, bloque tras bloque. Considerado de esta forma, la complejidad del estudio se simplifica enormemente, porque independientemente del tamaño que tenga la imagen, su análisis quedará reducido a considerar un número enorme de bloques, pero con un proceso repetitivo, que obliga a que el interés se centre únicamente en lo que se va a hacer con cada uno de estos bloques.

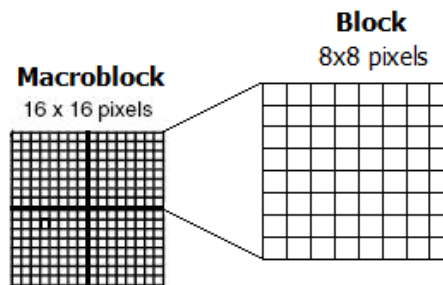


Figura 7-2. Descomposición en bloques de un macrobloque.

7.2.3 Señales utilizadas

De cada uno de los píxeles del macrobloque se obtienen tres informaciones diferentes, que son:

- Información de luminancia (señal Y).
- Información de una de las señales de cromaticidad (señal Cr).
- Información de la otra señal de cromaticidad (señal Cb).

En el caso de MPEG, estas tres señales no están ligadas forzosamente a las señales RGB a través de ecuaciones inamovibles, sino que dicha relación puede ser de cualquier tipo ya que se envían en cada caso particular al decodificador para que este sepa cuál es el significado de las señales que recibe y pueda obrar en consecuencia.

Para la señal de luminancia es necesario transmitir las informaciones de todos los píxeles, sin embargo, para las señales de crominancia, sabiendo que el ojo para ellas es menos exigente, es posible suprimir alguna de las muestras. Dependiendo del número y posición de las muestras de crominancia despreciadas, aparecen varios tipos de formas de muestreo, que se encuentran definidos por la variable **formato de crominancia** (**chroma_format**), transmitida en la cabecera de extensión de secuencia.

Esta variable enviada consta de dos bits, por lo que puede informar de cuatro situaciones diferentes, según se indica en la Tabla 7-2:

Tabla 7-2. Diferentes valores para la variable chroma_format.

Chroma_format	Significado
00	Reservado
01	4:2:0
10	4:2:2
11	4:4:4

La variable chroma_format recibida nunca valdrá 00. Esta combinación se suele eludir para todas las variables en la mayor parte de los casos, para evitar un error en la interpretación de las variables ya que podrían aparecer muchos ceros seguidos. Además en este caso solo se necesitan tres combinaciones para reflejar todos los casos posibles. Puede ser que en un futuro si sea necesario el uso de la combinación '00' por ello recibe el nombre de 'reservado' y no de 'prohibido'.

Formato de crominancia 4:4:4 (chroma_format = 11). Este tipo de formato indica que no se ha despreciado ninguna de las muestras de crominancia y que por lo tanto se utilizan todas. Para cada uno de los píxeles existirán tres señales, la de luma y las dos de croma. La adquisición de los datos será la indicada en la Figura 7-3.

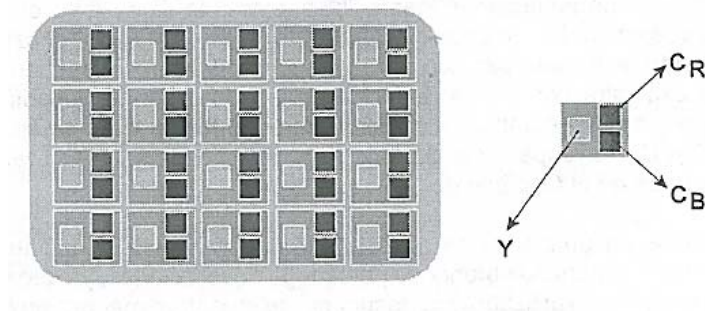


Figura 7-3. Muestras de luminancia y crominancia para 4:4:4.

Formato de crominancia 4:2:2 (chroma_format = 10). En este formato, tal como se indica en la Figura 7-4, sobre cada línea se desprecian la mitad de las muestras de crominancia, eligiendo una sí y otra no. El tratamiento es el mismo para todas las líneas, por lo que en sentido vertical, hay columnas que en cada píxel tienen 3 muestras (una de luma y dos de croma) y hay columnas en las que solo hay muestras de luminancia.

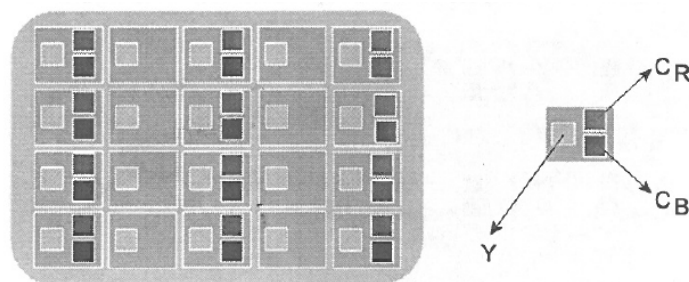


Figura 7-4. Muestras de luminancia y crominancia (formato 4:2:2).

Formato de crominancia 4:2:0 (chroma_format = 01). Cuando los datos se obtienen de esta forma, además de despreciar en sentido horizontal para las señales de crominancia una de cada dos muestras, también se desprecia una de cada dos muestras en sentido vertical. La situación en este caso es representada en la Figura 7-5.

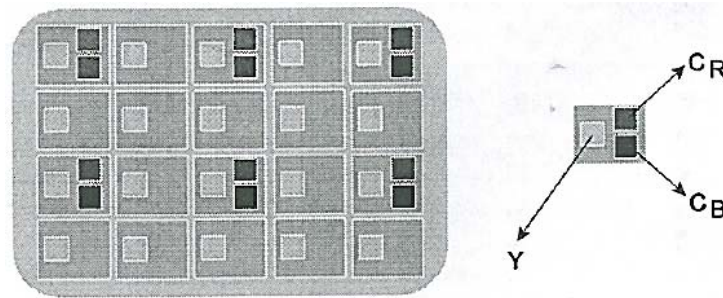


Figura 7-5. Muestras de luminancia y crominancia (formato 4:2:0).

Dentro de un macrobloque, los bloques deben numerarse, porque hay que tenerlos totalmente identificados para no confundir unos con otros y entre otras cosas conocer el orden en el que serán enviados. En la Figura 7-6 se indica la numeración asignada a los bloques para los tres formatos de muestreo posibles.

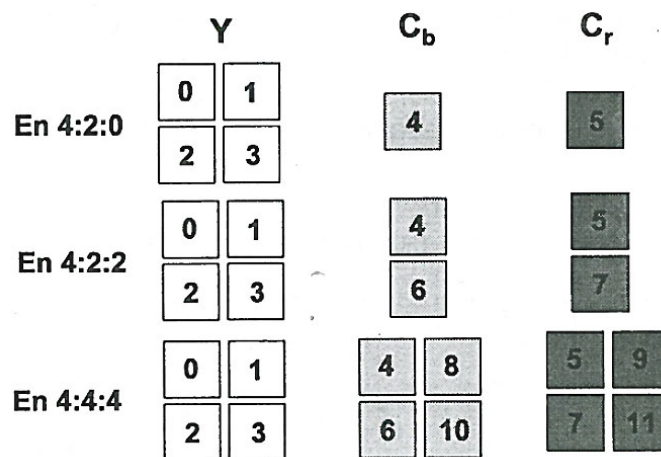


Figura 7-6. Numeración de los bloques dentro del macrobloque.

7.2.4 Transformada discreta del coseno

Una transformada es una operación matemática que permite pasar a un dominio transformado, en el que se resaltan algunos detalles de forma bastante más precisa que en el dominio original.

La transformada utilizada Discreta del Coseno (DCT) es una transformada bidimensional. Para efectuar esta transformada se toma el bloque $x(m, n)$, que serán valores de brillo si es bloque de luminancia o en otro caso valores de crominancia y arrojará a su salida otro conjunto de 8×8 valores denominados coeficientes. Esta transformada es reversible, por lo que a partir de los coeficientes se pueden obtener de nuevo los valores originales en el momento que se desee. A esta operación se la conoce como transformada inversa (IDCT).

Puede decirse que los 64 coeficientes obtenidos al hacer la transformación de un bloque, son las cantidades que hay que utilizar de las 64 imágenes que aparecen en Figura 7-7 denominadas imágenes base para que al mezclarlas se obtenga la imagen original.

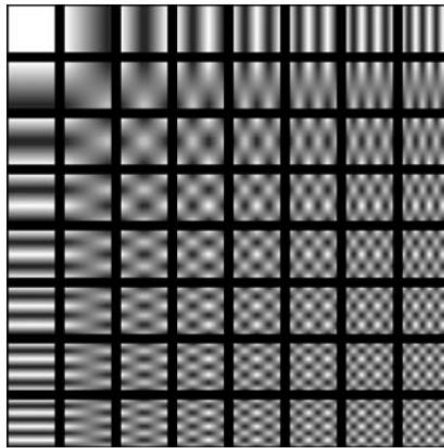


Figura 7-7. Conjunto de imágenes base.

Después ya se entiende que si los 64 valores del bloque a transformar son iguales, de los 64 coeficientes obtenidos solo uno de ellos será distinto de cero, y los otros 63 serán nulos. Esa es la gran ventaja de esta transformada, que obtiene una gran cantidad de ceros, lo que será muy fácil de codificar posteriormente.

7.2.5 Cuantificación

Los valores obtenidos al aplicar la DCT a un bloque pueden tener una precisión muy grande, todo lo que se quiera, por lo que si se les codificase directamente se necesitaría un número de bits muy grande para identificar de forma binaria a cada uno de los coeficientes.

Por eso, se procede a una cuantificación, utilizando para ello un determinado escalón cuántico, lográndolo mediante la expresión (7-1):

$$\text{índice} = \text{redondeo de } \left\lfloor \frac{\text{Coeficiente}}{\text{Escalón cuántico}} \right\rfloor. \quad (7-1)$$

Y de esta manera se pasa de coeficientes a coeficientes codificados, siendo normal referirse ellos con el nombre de Índices.

De los 64 coeficientes, el primero de ellos se denomina coeficiente DC y los otros 63 reciben el nombre de coeficientes AC. El primero de ellos expresa el valor medio del brillo del bloque, pudiendo tomar cualquier valor, mientras que los demás expresan las variaciones que existen dentro de él, por lo que suelen ser valores muy pequeños. En la Figura 7-8 se muestra esquemáticamente el curso que sigue un bloque hasta que llega al punto de la cuantificación.

Precisamente debido a esta diferencia, también sufren posteriormente un proceso de codificación diferente. El coeficiente DC de cada bloque se codifica diferencialmente respecto al valor del bloque anterior, con objeto de codificar valores que casi siempre sean pequeños y necesiten pocos bits para identificarlos. Los coeficientes AC (que ya son siempre pequeños de por sí) se codifican directamente.

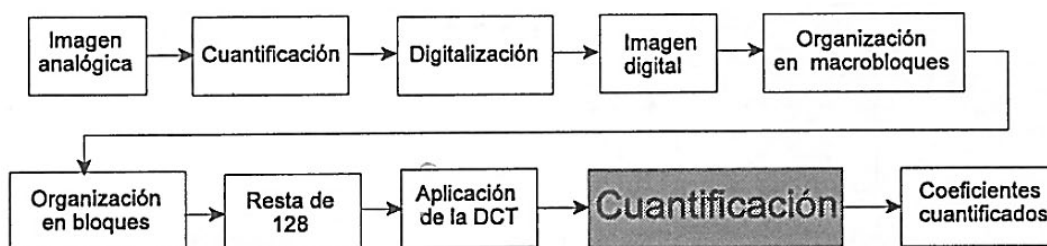


Figura 7-8. Bloque de cuantificación que se está estudiando.

Ya se ha dicho anteriormente que de los 64 coeficientes, el único que sufre un tratamiento diferente es el primero, el que está colocado en la posición (0,0) y que como proviene de una transformación frecuencial a través de la Transformada Discreta del Coseno, se denomina coeficiente de corriente continua o coeficiente DC.

Si el bloque es codificado tipo 'Intra', no se envía el valor absoluto, sino la diferencia respecto a su homólogo del bloque anterior en la misma imagen. En la Figura 7-9 se intenta aclarar gráficamente este concepto, pudiendo apreciarse que en ese caso particular de este ejemplo, se deberá enviar el valor $(28 - 26) = +2$.

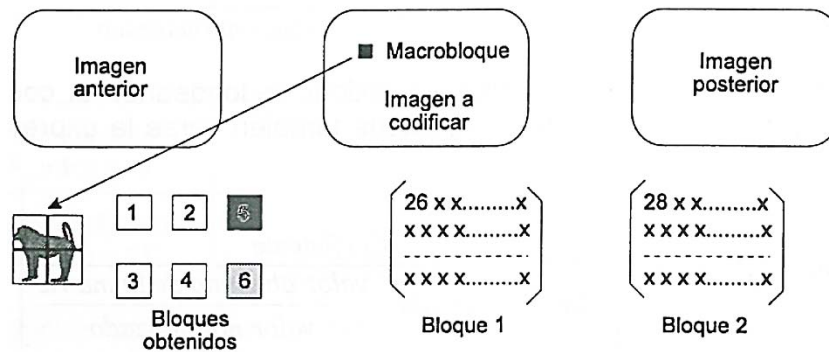


Figura 7-9. Valores implicados en la codificación diferencial.

En el decodificador, a la llegada del índice, se puede volver a recuperar el valor del coeficiente con la expresión (7-2):

$$\text{Coeficiente recuperado} = \text{Índice} * \text{Escalón cuántico. (7-2)}$$

El valor del escalón cuántico se envía camuflado en dos magnitudes para facilitar su transporte.

Este factor de escala es el mismo para los 64 coeficientes del bloque. Hay una única excepción y es que para el coeficiente DC de los bloques codificados tipo intra se sigue un tratamiento diferente, pero solo en ese caso.

Se puede especificar para cada uno de los formatos de muestreo las matrices de pesos que van asociadas a los bloques Intra y a los bloques no Intra. Si se decide especificar concretamente las matrices que han utilizado en la cuantificación y por lo tanto las que deberán ser utilizadas en la decuantificación, será necesario enviarlas o bien en la cabecera de secuencia o en una extensión específica. Este es el procedimiento genérico y es el que servirá para los 63 coeficientes AC, cuando el bloque sea Intra codificado o para los 64 coeficientes cuando sea no Intra codificado.

7.2.6 Codificación tipo Intra

En una sucesión de imágenes, normalmente existe mucha redundancia entre una imagen y las siguientes, por lo que a menudo se recurre a la técnica de no enviar completamente una imagen, sino enviar solo la diferencia con otras imágenes, que ya se encuentran disponibles en el decodificador, porque han sido enviadas con anterioridad. Esta técnica ahorra enormemente el consumo de bits de información.

Sin embargo, cada cierto tiempo es necesario enviar una imagen completamente, conociendo y olvidando el hecho de que el decodificador es posible que posea informaciones que van a ser enviadas de nuevo. A continuación se resume por qué esto es necesario:

Los decodificadores pertenecen al público en general y por lo tanto serán conectados cuando cada usuario lo desee. Hasta que no llegue una de estas imágenes completas no podrán comenzar a funcionar, ya que no disponen de las informaciones anteriores.

Si se produce un fallo en la recepción de señales, mientras no llegue una de estas imágenes completas se cometerán gran cantidad de errores. Desde el momento en que se recibe el funcionamiento será correcto.

Cuando se procede a editar informaciones diferentes para confeccionar un programa, el montador ordena la conmutación de un conjunto a otro cuando su gusto artístico se lo aconseje.

Existen otras múltiples razones que aconsejan que con una frecuencia relativamente alta, se codifique una imagen de forma completa, olvidándose de las informaciones que le han precedido.

Cuando se pretende codificar una imagen utilizando este tipo de codificación, únicamente se tiene en cuenta la información de la propia imagen que se está considerando, sin que importen en absoluto las informaciones de imágenes anteriores o posteriores.

Dado que una imagen tendrá bastantes píxeles similares no parece muy lógico enviar píxel tras píxel y para solucionar esto se implementan las imágenes codificadas según lo que se denomina tipo Intra (o tipo I en inglés *Intra Frame Coding*).

Para un conjunto de 16x16 píxeles (macrobloque) y suponiendo que los datos de luminancia y de crominancia se extraen según la estructura de muestreo 4:2:0, el diagrama de bloques utilizado para la codificación de dicho conjunto mediante la técnica tipo I, es el representado en la Figura 7-10.

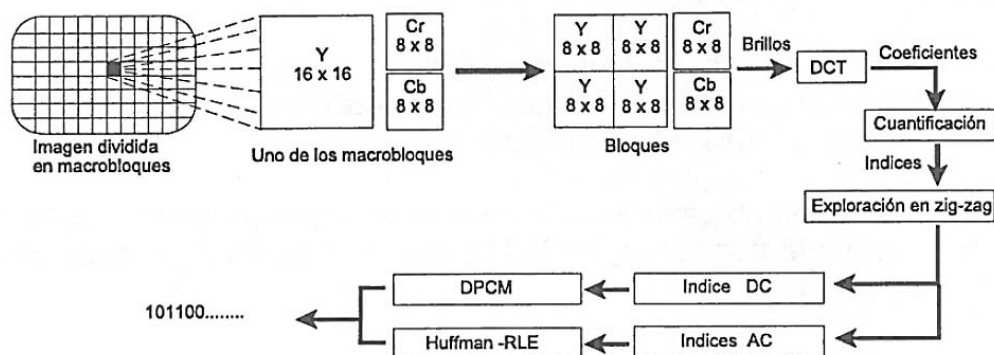


Figura 7-10. Codificación de una imagen tipo Intra.

De cada macrobloque se extraen un número de bloques (12, 8 o 6 dependiendo del formato de muestreo elegido). Sobre cada uno de ellos, formado por un conjunto de 64 valores (una matriz), se llevan a cabo un conjunto de operaciones que se describen a continuación.

Las mismas operaciones, que de forma esquematizada aparecen en la Figura 7-11, que se hacen para un bloque se hacen para todos los demás.

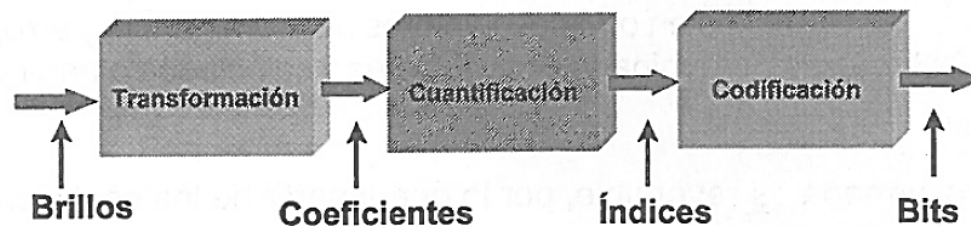


Figura 7-11. Proceso básico para un bloque.

7.2.6.1 Codificación del coeficiente DC en bloques tipo Intra

El proceso es el siguiente:

- El coeficiente se cuantifica según se especificó en el apartado anterior.

- Se codifica en DPCM, calculando la diferencia con el índice para la misma posición DC del bloque anterior, salvo en el caso del primer bloque, que se comparará con el valor central. Ese valor, que normalmente será bastante pequeño, es el que se transmitirá.

Para la técnica de muestreo 4:2:0, de los 6 bloques que forman el macrobloque se numeran a los bloques tal como se establece en la Figura 7-12 (0, 1, 2, 3 para luma y 4, 5 para cromina). En el se explica la obtención de los códigos de los bloques de luma y cromina.

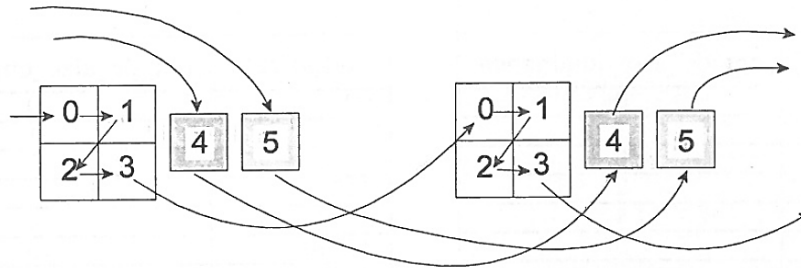


Figura 7-12. Codificación diferencial de los coeficientes DC.

Exploración en zig-zag

Una vez cuantificados de la forma descrita a los 64 coeficientes, teniendo en cuenta que se encuentran distribuidos en forma de matriz, es necesario formar con ellos una secuencia unidimensional, colocando unos a continuación de los otros.

Esto se hace mediante la exploración conocida como zig-zag y dicha exploración puede ser de dos tipos, dependiendo de cuanto valga una variable que también acompaña a la señal y que recibe el nombre de `alternate_scan`, que como solo adopta dos posiciones diferentes es suficiente con un bit para identificarla.

En la Figura 7-13 se representan las dos posibles formas de exploración en zig-zag.

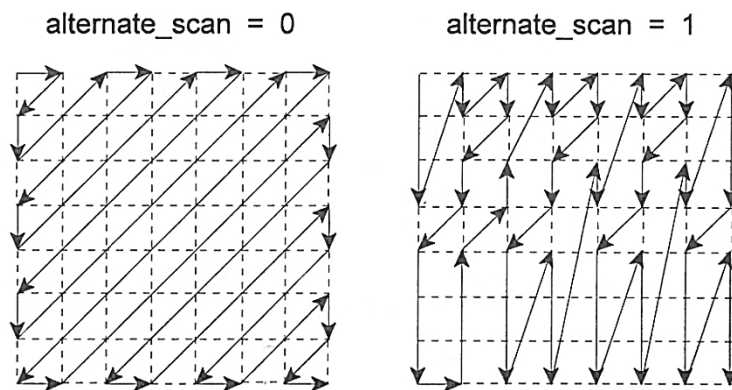


Figura 7-13. Dos posibles formas de exploración en zig-zag.

Dadas las características de la matriz de índices, al final de la secuencia completa de **62 índices AC aparecerán** muchos valores nulos. Obviamente esos valores no serán codificados, sino que cuando en la secuencia de 63 índices aparezca el último valor significativo (distinto de cero), ya no se siguen codificando los ceros que vienen detrás, sino que se coloca un código conocido como Fin de Bloque, que será interpretado precisamente como eso, como todos los índices que aún no se han decodificado pasan a tomar inmediatamente el valor cero. En el **¡Error! No se encuentra el origen de la referencia.** se detallan las tablas utilizadas para la codificación de los índices AC en los bloques tipo Intra.

7.2.6.2 Codificación de los índices AC en bloques tipo Intra

Estando inmersos en la codificación de macrobloques tipo I, a partir de la secuencia unidimensional de índices obtenida en el apartado anterior, es necesario proceder a su codificación.

Para su codificación se acude a una Tabla en la que mediante un método Huffmann y en función de las probabilidades de aparición se han establecido los códigos correspondientes.

7.2.7 Codificación tipo Inter

Cuando una imagen se codifica de acuerdo a este criterio, la información que se envía no es relativa a la imagen en sí misma, sino que la información codificada es relativa a la diferencia que existe con imágenes anteriores, con imágenes posteriores y en alguno de los casos, tanto con imágenes anteriores como posteriores.

En MPEG-1 solo hay un tipo posible de codificación tipo inter, pero en MPEG-2 existen dos formas posibles de codificación tipo inter, también denominada a veces tipo no intra, para distinguirla del anterior tipo codificación. Estas dos formas de codificación dan orígenes a imágenes que se las denomina tipo P o tipo B.

En el diagrama de bloques de la Figura 7-14 se da a entender que el decodificador, antes de que le llegue la imagen en cuestión, en función de otros datos recibidos anteriormente, efectúa una predicción de lo que le va a llegar.

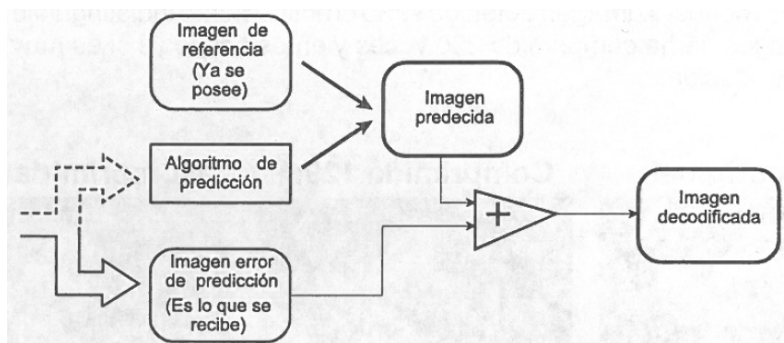


Figura 7-14. Obtención de la imagen en el codificador.

Esta predicción se efectúa en base a un algoritmo que se denomina algoritmo de predicción y que en cada aplicación será diferente.

El codificador no envía la imagen, sino que como es conocedor de la predicción que realizará el decodificador, deberá enviar únicamente aquella información que añadida a la que ha predicho, proporcione la imagen real. Si el algoritmo de predicción ha sido bastante bueno, la información adicional no deberá ser muy grande, y con muy poco gasto de transporte se podrá disponer en el decodificador de la imagen completamente perfecta, sin ningún deterioro en el proceso.

Este proceso se encuentra reflejado en el diagrama de bloques de la Figura 7-15.

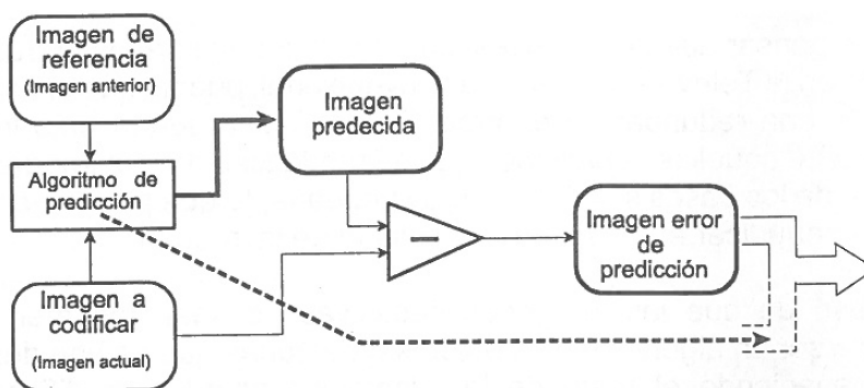


Figura 7-15. Obtención de las informaciones a enviar por el codificador.

En la mayoría de los casos las imágenes no son idénticas, pero tienen la característica de que hay muchos objetos que son comunes, aunque hayan cambiado de posición.

Para poder comprender bien el algoritmo de predicción utilizado y en consecuencia la imagen de error que será necesario enviar, es necesario analizar primero, aunque sea brevemente, un concepto que se denomina vector de movimiento.

7.2.7.1 Vector de movimiento

Al codificar una imagen, se puede comprobar que existe mucha redundancia espacial, dado que un determinado píxel tiene mucha probabilidad de ser idéntico a los píxeles que se encuentran en su proximidad.

Esa redundancia espacial se aprovecha para comprimir el tamaño de la imagen y lograr reducirla a un tamaño muchísimo menor que el que tenía originalmente, sacrificando poco o nada la calidad de la misma. Eso se consigue mediante la Transforma DCT, la cuantificación y codificación mediante códigos de longitud variable.

Sin embargo, existe otro tipo de redundancia entre las imágenes que componen una secuencia de vídeo y es que entre ellas habrá muy poca diferencia de contenido a no ser que se produzca un cambio de plano. Esto hace pensar que no es necesario transmitir una imagen detrás de otra, como se venía haciendo en la televisión analógica, sino solamente la diferencia o variación que se produce de una imagen a otra, que en la mayoría de casos será nula o muy pequeña, lo que proporciona un enorme ahorro de bits.

En estas condiciones resulta obvio que no es necesario transmitir el fondo de la escena sino solo el desplazamiento que han sufrido determinados objetos de la escena.

En el ejemplo de la Figura 7-16 de manera simplificada se viene a explicar el concepto de vector de movimiento, considerando que solo se mueve el balón.

**Figura 7-16. Vector de movimiento para el balón.**

Por tanto, para cada uno de los macrobloques, se busca en una zona alrededor de esa posición, en la imagen que va a ser su referencia, a ver si se encuentra un trocito del mismo tamaño que el macrobloque, con el que haya una gran coincidencia y del resultado de esa búsqueda pueden suceder cuatro cosas:

- Que se encuentre ese trocito de información que coincide completamente con el macrobloque que se va a codificar en su misma posición. En este caso el decodificador se salta dicho macrobloque (*skipped*) ya que con los datos que tiene en la memoria puede recomponerlo sin problemas.
- Que se encuentre ese trocito de información que coincide completamente con el macrobloque que se va a codificar pero en una posición ligeramente desplazada. En este caso en la cabecera del macrobloque se indica lo que hay que desplazar ese trocito de información para colocarlo adecuadamente en dicho macrobloque.
- Que se encuentre ese trocito de información que coincide solamente en parte con el macrobloque que se va a codificar, en cuyo caso se tendrá que enviar la diferencia existente entre el macrobloque de la

imagen anterior y el del macrobloque a codificar más el vector en el caso de que también estuviese desplazado.

- Que en la zona de búsqueda se encuentre en la imagen de referencia un trocito de información completamente diferente a la del macrobloque que se va a codificar. En este caso, este macrobloque deberá llevar una codificación tipo 'intra', luego no hay vectores de movimiento en la cabecera del macrobloque.

El proceso de codificación-decodificación no es simétrico, algo que se intenta conseguir en todos los sistemas de transmisión en los que hay un solo codificador y muchos posibles decodificadores.

A la vista de todos estos conceptos ya se intuye que el algoritmo de predicción no es simplemente tomar la imagen de referencia, sino que se toma como predicción la imagen que se obtiene de considerar la imagen de referencia, pero también los vectores de movimiento, puesto que esta predicción será mucho más cercana a la realidad.

7.2.7.2 Imágenes tipo P

En un razonamiento en principio, sin entrar en gran profundidad, se comprende que enviar una imagen, cuando se ha enviado antes otra que tiene prácticamente la misma información, es algo que parece que no tiene mucho sentido, si se pretende ahorrar bits en la transmisión. En efecto, así es la situación y eso es lo que intenta conseguir el establecimiento de imágenes tipo P. Aprovechar todo lo que se pueda de lo que ya se ha enviado y por lo tanto está disponible en el decodificador sin necesidad de enviarlo de nuevo.

Cuando una imagen se codifica de esta forma, para conocerla perfectamente solo hay que tener en cuenta los datos que se reciben en la propia imagen y los de otra imagen que la precedió y que ya ha sido recibida anteriormente.

Los datos recibidos en la propia imagen solamente albergan la información diferencia entre la imagen que se está procesando en ese momento (en inglés se denomina *target image* y aquí se denominará imagen actual o imagen a codificar) y otra imagen anterior que solo será una, que se denomina imagen de referencia (*reference image*), después de haberse desplazado lo indicado por los vectores de movimiento. Lógicamente, también hay que enviar los referenciados vectores de movimiento.

En la Figura 7-17 se puede apreciar como a partir de la imagen de referencia y los vectores de movimiento se obtiene una imagen que forma justo la predicción que el decodificador va a imaginar sobre la imagen que le va a llegar a continuación. La imagen anterior ya la tiene decodificada y completamente almacenada y los vectores de movimiento llegan en la propia imagen, pero también se dispone de ellos.

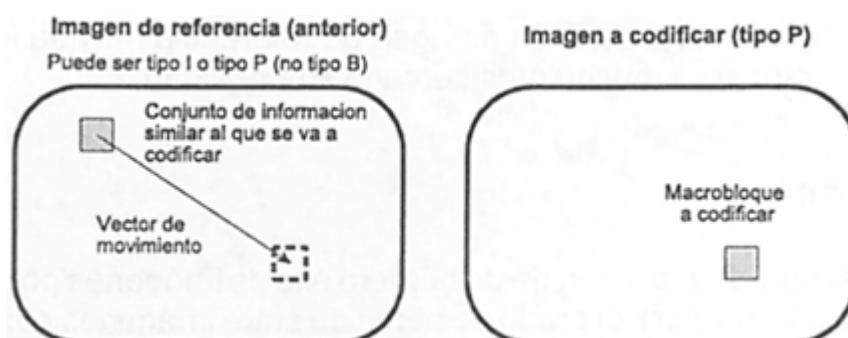


Figura 7-17. Codificación de una imagen tipo P.

Como contenido de la propia imagen llegará la diferencia existente entre la imagen real y la que se ha obtenido en la predicción explicada en el párrafo anterior, por lo que sumándola a la predicción, se conseguirá disponer perfectamente de la imagen original.

La imagen de referencia solo puede ser de tipo I o de tipo P, pero nunca podrá ser de tipo B.

Es importante notar que la imagen de referencia será la imagen, de uno cualquiera de estos dos tipos (I o P), que se encuentre más cercana a ella, siempre por la parte de atrás, o lo que es lo mismo, que se haya generado con anterioridad a ella.

En la Figura 7-18 se ha representado un conjunto de 13 imágenes, indicándose para cada una de ellas si es tipo I, tipo P o tipo B. En este ejemplo, para la imagen 4 se tomará como imagen de referencia a la imagen 1, ya que es la imagen más cercana por la parte de atrás que puede ser utilizada para este fin. En este caso, la imagen de referencia es tipo I.

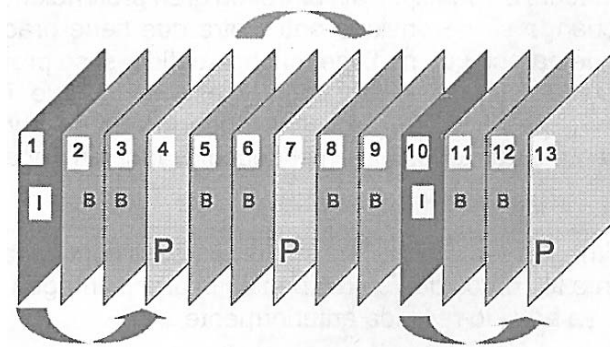


Figura 7-18. Ejemplo de un conjunto de imágenes.

La letra P con la que se identifica a este tipo de codificación, está extraída de la inicial de *Predictive-codec Picture*, puesto que realmente lo que el codificador envía es lo que necesita el decodificador para añadirse a lo que de forma predictiva él ha calculado para esa imagen. Así pues, en una imagen codificada tipo P, dentro de cada uno de los macrobloques pueden ir albergados vectores de movimiento, con objeto de que no sea necesario incluir su contenido, ya que puede irse a buscarlo a una imagen anterior, en la posición indicada por el vector de movimiento.

La codificación de la diferencia existente entre el macrobloque a transmitir y el trozo de imagen (no macrobloque) que se va a utilizar de la imagen anterior debe hacerse en base al diagrama de bloques indicado en la Figura 7-19.

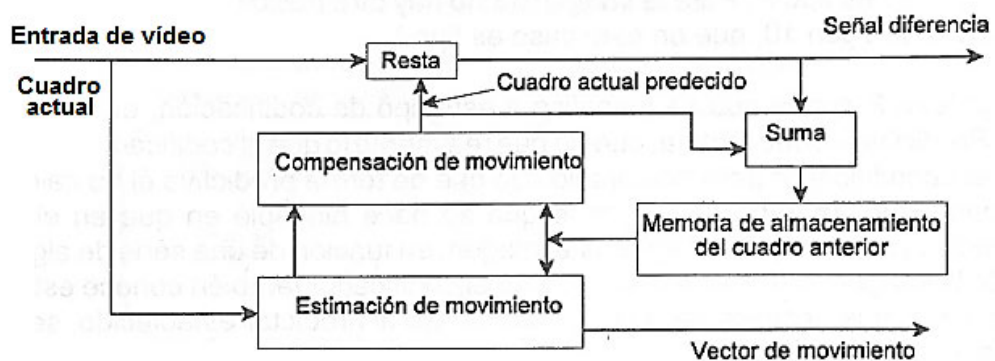


Figura 7-19. Codificación tipo P

Los cuadros codificados tipo P consiguen mucha más compresión que los cuadros codificados tipo I, ya que se apoyan en informaciones transmitidas anteriormente y que por lo tanto no hay que transmitir de nuevo, pudiéndose estimar que el número de bits que conforman un cuadro tipo P es del orden de un tercio de los bits que hay normalmente en un cuadro tipo I. Puede establecerse que para las imágenes codificadas tipo P se suelen conseguir compresiones del orden de 0,2 bits por píxel.

Debe quedar muy claro que dentro de una imagen codificada tipo P pueden existir macrobloques para los que no se haya podido encontrar en la imagen de referencia un conjunto de información muy similar, por lo que habrá

que recurrir a codificarlo tipo 'intra'. Esto se puede prestar a confusión para los que no denominen perfectamente todas estas técnicas, y por esto se insiste en que debe ser contemplada esta posibilidad, siendo por ello que cada macrobloque de alguna forma indica el tipo de codificación que ha sufrido, para que el decodificador sepa a qué tipo de macrobloque se está enfrentando.

Como resumen final de este tipo de imagen tipo P, pueden establecerse las posibilidades que se pueden presentar al intentar codificar un determinado macrobloque, que ya han sido comentadas con anterioridad y que pueden resumirse en las siguientes:

- Que el macrobloque a codificar coincida en información con la que se encuentra, completamente idéntica y en la misma posición en la imagen de referencia. Esta situación será muy frecuente y cuando sucede, simplemente se salta el macrobloque. El ahorro en este caso es total.
- Que la información del macrobloque se encuentre idéntica, pero en otra posición diferente a la ocupada en la imagen a codificar. En este caso solo se enviará el vector de movimiento. El ahorro es enorme.
- Que se encuentre parecida la información en la imagen de referencia, pudiendo estar en la misma o en cercana posición. En este caso se envía el vector de movimiento (si es que existe) y la diferencia entre esas informaciones. El ahorro de bits es importante, menos que en el caso anterior.
- Que aunque se busque esa información en la imagen de referencia no se encuentre, ni igual, ni parecida. En ese caso no hay más remedio que codificar el macrobloque como si se tratase de un macrobloque de una imagen tipo I. Es el caso peor que puede producirse, pero incluso en este caso se aprovechará la redundancia espacial existente dentro del macrobloque y algo se comprimirá.

7.2.7.3 Imágenes tipo B

Además de las imágenes tipo I (que únicamente contienen información sobre sí mismo) y las imágenes tipo p (que contienen información a añadir sobre la predicción hecha sobre una imagen anterior) existen otro tipo de imágenes, denominadas tipo B.

Las imágenes codificadas tipo P son muy efectivas si no existe un cambio de plano entre la imagen que se pretende codificar y la imagen que sirve como referencia. En el caso de que existe un cambio de plano entre estas dos imágenes, no se obtiene mucho rendimiento en este tipo de codificación.

Cuando una imagen se decide codificar tipo B, se piensa aprovechar tanto lo que se puede de una imagen anterior, como lo que se pueda de una imagen posterior. Actuando de esta forma se asegura que alguna de las imágenes estará relacionada con la que se pretende codificar, sin que exista un cambio de plano entre ellas. Si existe este cambio de plano con la imagen anterior es muy improbable que lo exista con la imagen posterior o viceversa. La situación normal será la que puede apreciarse sobre la Figura 7-20.

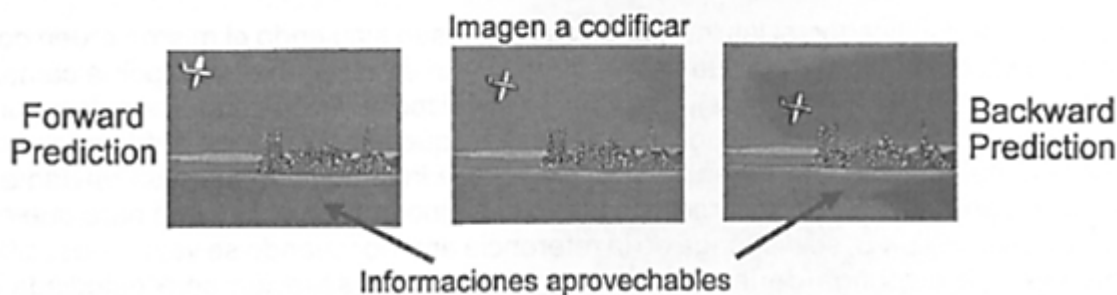


Figura 7-20. Posibilidades de las imágenes tipo B

En esta forma de codificación, solo se aporta información complementaria sobre la que se puede obtener a partir de disponer completamente de dos imágenes, una anterior y otra posterior, que pueden ser de tipo I o de tipo P.

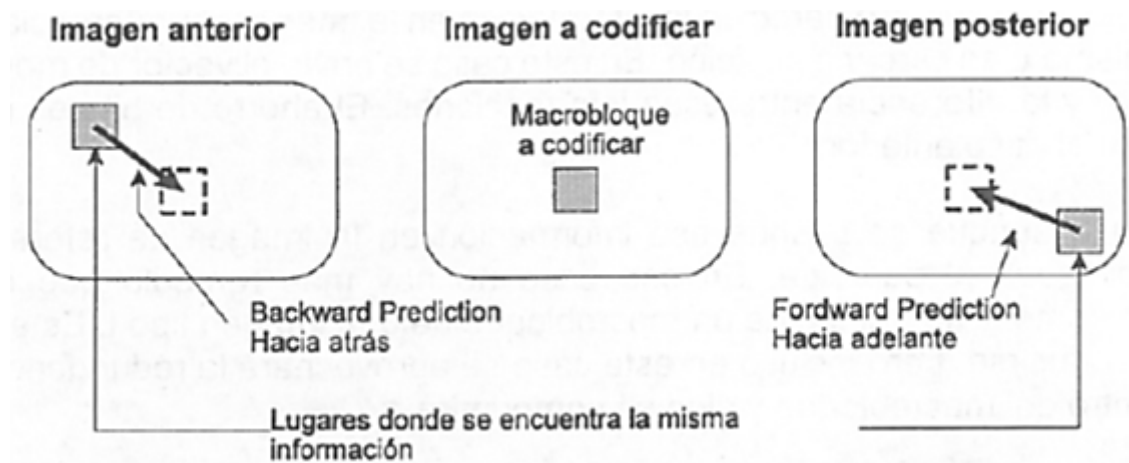


Figura 7-21. Predicciones para las imágenes codificadas tipo B

Para las imágenes que se codifican tipo B, se dispone de una imagen de referencia anterior (*backward-predicted image*), pero también se dispone de otra imagen de referencia posterior (*forward-predicted image*). La información que va en la imagen que se va a codificar tipo B, es precisamente aquella que falta para determinar completamente la imagen, a partir de las dos imágenes de referencia y los correspondientes vectores de movimiento. Este tipo de codificación queda reflejado claramente en la Figura 7-21.

En principio puede pensarse que tanto en el decodificador como en el codificador es lógico disponer de la imagen anterior, pero puede parecer que es imposible disponer de la imagen posterior, cuando se está tratando de codificar o decodificar la imagen actual.

En el codificador esto se soluciona retardando un poco toda la información, por lo que no se codifica la imagen actual hasta que no se disponga de su referencia posterior. Una vez conocida, se codifica de acuerdo a ella y aunque se ha retenido la información hasta que llegue la imagen siguiente, el retardo es tan pequeño que no origina grandes problemas, solo el de que es necesario disponer de cantidades enormes de memoria con características de elevada velocidad de escritura y lectura.

En el decodificador, si las imágenes se recibiesen siguiendo el mismo orden con el que acceden al codificador, sería imposible disponer de la imagen siguiente antes que la que se pretende decodificar, pero la realidad es que aun a costa de sufrir más retraso en toda la señal, se altera el orden de transmisión de las imágenes y no se envían en el mismo orden en que se han capturado ni codificado, sino que se acomodan para que en el caso de un cuadro tipo B, se disponga de la referencia anterior cuando se vaya a decodificar, pero también se disponga de la referencia siguientes. Esta situación será estudiada con detalle cuando se describa pormenorizadamente el orden de emisión de las imágenes en una secuencia de vídeo.

La denominación de cuadro tipo B proviene de la inicial de la inicial de *Bidirectionally predictive-code picture*, que es como se define en inglés.

Con objeto de aclarar los conceptos, en la Figura 7-22 se representa una sucesión de imágenes, indicándose cuáles son las imágenes de referencia que se utilizan para decodificar cada una de ellas. Puede apreciarse que para decodificar la imagen 2 se hace a partir de los que llega en la propia imagen más la imagen 1 y la imagen 4, que deberán decodificarse antes.

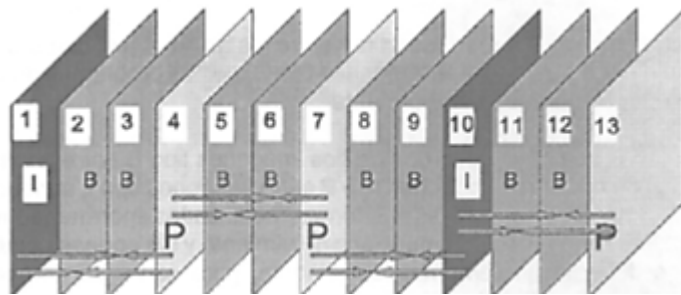


Figura 7-22. Imágenes utilizadas para obtener las predicciones.

La imagen 3 necesita también de la 1 y de la 4. Visto así, puede parecer que la imagen 3 coincidirá con la 2, pero eso no es así porque aunque tenga las mismas imágenes de referencia, su propio contenido es diferente.

Las imágenes tipo B llevan en sí mismas muy poca información, puesto que el predictor que se hace a partir de dos imágenes, una que se encuentra delante y otra que se encuentra detrás y sus correspondientes vectores de movimiento, está muy cercano a la imagen real, lo que indica que no será necesario apenas transmitir información.

La distancia entre imágenes tipo I viene condicionada por aspectos operativos tales como el hecho de que cuando se editan y se mezclan imágenes que llegan de diferentes lugares, hay que esperar a que llegue una imagen tipo I, porque son las que reinician el sistema y lo dejan configurado independientemente de las informaciones anteriores. Si esta distancia fuese muy grande, los puntos de corte se encontrarían fuertemente limitados al estar muy distanciados entre sí.

Las imágenes tipo P deben repartirse entre las de tipo I y el número de ellas debe ser tal que no deje entre ellas muchas imágenes tipo B dado que para almacenarlas se necesita mucha memoria y por lo tanto buffers muy grandes.

Se deduce por tanto que la elección de utilizar imágenes tipo I, P o B no se hace condicionada por el contenido de las imágenes, sino que la decisión es totalmente técnica.

Mediante dos parámetros queda definida totalmente la distribución relativa de imágenes. Estos dos parámetros se denominan 'n' y 'm' y sobre la Figura 7-23 se puede apreciar su significado. 'n' establece la distancia entre imágenes tipo I y dentro de ellas, el valor 'm' establece la distancia entre imágenes tipo P. Un valor típico es el de 15/3 como se muestra en la Figura 7-24.

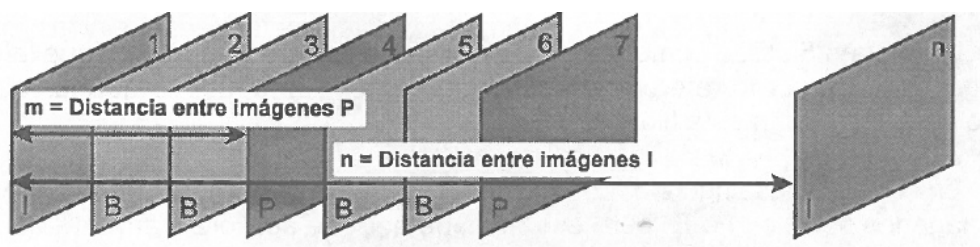


Figura 7-23. Establecimiento del ritmo de imágenes I, P y B.

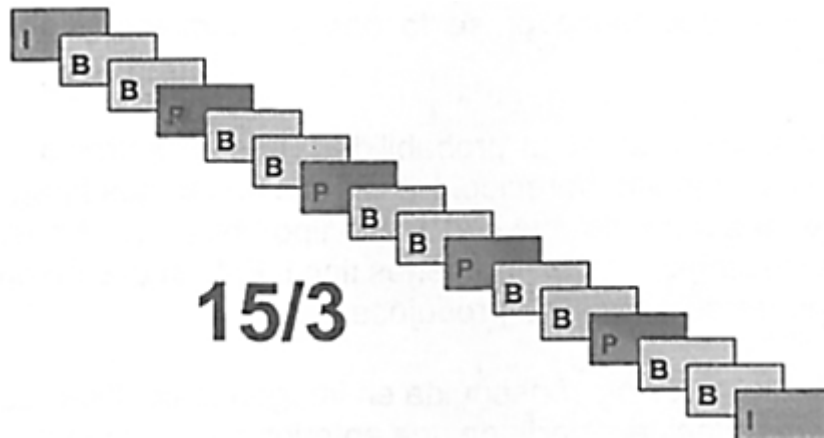


Figura 7-24. Estructura típica de distribución de imágenes.

En la codificación tipo B, a la hora de decidir codificar un macrobloque, se pueden presentar varias situaciones, que es necesario comentar por separado cada una de ellas, debido a la gran importancia que posee este aspecto.

- Que dicho macrobloque coincida completamente en posición con las dos imágenes de referencia. En este caso, simplemente se salta ese macrobloque y no se le presta atención en la codificación. El decodificador ya sabe que puede acudir a cualquiera de las imágenes de referencia y extraerlo de ellas.
- Que dicho macrobloque coincida completamente en una de las imágenes de referencia. Si eso es así, el codificador únicamente deberá informar en cuál de las imágenes de referencia se encuentra, para que el decodificador sepa dónde debe ir para localizar esa información.
- Que dicho macrobloque se encuentre parecido en las dos imágenes de referencia. Esta es una situación muy frecuente y en este caso se actuará de la siguiente forma:

Entre las dos informaciones ‘parecidas’, tanto en el codificador como en el decodificador, se hace la media entre ellas y ese resultado es muy probable que coincida con el macrobloque que se pretende transmitir. Esa es la primera predicción establecida y lo que hay que enviar es simplemente los vectores de movimiento (si los hay) y la diferencia respecto a esta predicción, que lógicamente será muy pequeña. Con las dos imágenes decodificadas y los vectores de movimiento es con lo que se obtiene la predicción definitiva.

- Que dicho macrobloque se encuentre bastante parecido en una de las imágenes de referencia. En este caso se envía el vector de movimiento y la diferencia existente entre el macrobloque que se pretende enviar y el trozo de imagen ‘parecido’ que se ha encontrado. Lógicamente, hay que enviar la información de cual las dos imágenes es la que interesa.

El grado ‘parecido’ que se nombra es elegido por el propio codificador.

- Por último y con muy poca probabilidad, puede ocurrir que el macrobloque a codificar no se encuentre igual ni parecido en ninguna de las dos imágenes de referencia. En este caso no hay más remedio que codificarlo tipo ‘intra’ de la misma manera que se hacía para todos los macrobloques en imágenes tipo I. Es una situación que apenas se producirá, pero hay que estar preparado por si se produjese.

Es enorme la compresión conseguida en imágenes codificadas según este tipo B, puesto que la imagen predeterminada a partir de una anterior y una siguiente es casi coincidente con la imagen intermedia que se está codificando, por lo que la información que hay que añadir para identificarla completamente es muy pequeña. Se puede llegar a un nivel de compresión del orden de 0,1 bits por píxel.

7.2.8 Estructura jerárquica de la señal

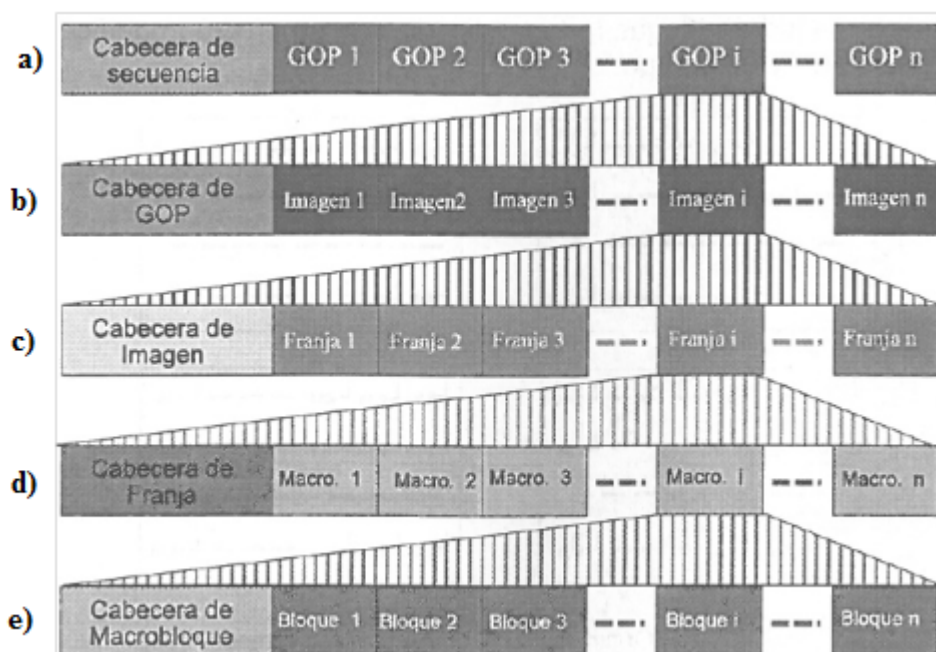


Figura 7-25. Estructura jerárquica de la señal de vídeo digital MPEG.

Para informar al decodificador del proceso que debe seguir, se deberán enviar gran cantidad de variables para indicárselo. Estas variables a veces afectan a muchas imágenes, otras veces afectan a pocas y en consecuencia es necesaria una gran organización para conocer en cada momento cuales son las variables que afectan en cada caso.

El conjunto más grande de información que puede encontrarse en MPEG se denomina secuencia. Este concepto no tiene nada que ver con el mundo artístico. Aquí no existe ningún criterio para establecer el comienzo de una secuencia, simplemente se comienza en el momento en el que se quiera. Una secuencia termina cuando se decide que va a comenzar otra nueva secuencia y no tiene un tamaño definido, se deja a criterio del codificador.

Cada secuencia lleva una cabecera, en la que se especifican parámetros que llevan toda la información de la misma: tamaño horizontal y vertical de las imágenes, relación de aspecto, número de imágenes por segundo, velocidad binaria de trabajo, matrices de cuantificación a utilizar, perfil y nivel con el que se va a trabajar, formato de muestreo, etc. La cabecera siempre arranca con un código que nunca aparecerá en el resto de la información, por lo que es totalmente imposible que no se interprete correctamente: 000001B_H.

La secuencia está formada por varios conjuntos de imágenes que se denominan grupos de imágenes, o normalmente GOP (*Group Of Pictures*). El número de ellos es ilimitado e impreciso, porque seguirán unos grupos detrás de otros, hasta que finalice la secuencia y ya se ha dicho que su tamaño no viene indicado ni siquiera en la cabecera de secuencia. La primera rama establecida en la estructura jerarquizada de organización de la información, es la que puede apreciarse en la Figura 7-25a.

Se entiende como grupo (GOP) al conjunto de imágenes que comienzan por una imagen tipo I y que por lo tanto no necesitan ayudarse para su decodificación de ninguna imagen del grupo anterior. En la Figura 7-25b se puede apreciar el desdoblamiento de un grupo, dando lugar a varias imágenes.

Las imágenes no son las unidades elementales de envío de información, sino que todavía se puede bajar más en la estructura jerarquizada. Por debajo de ella se encuentra el *slice* o franja. En la Figura 7-25c puede apreciarse este nuevo escalón jerárquico en la estructura de la señal MPEG, apreciándose como las imágenes se configuran en base a un conjunto de franja.

La franja tiene siempre una altura de 16 píxeles, pero la longitud no está fijada, viene indicada en su propia cabecera, suponiendo que esta exista, lo cual no es forzosamente obligatorio. Lo que sí es cierto es que la anchura es un múltiplo de 16 píxeles. La anchura de las franjas es de 16 píxeles porque esa es la altura del macrobloque en el que posteriormente se dividirán las franjas y cada franja deberá tener un número entero de macrobloques. Los dos últimos escalones jerárquicos que son el macrobloque y el bloque, vienen representados en la Figura 7-25d y Figura 7-25e respectivamente.

7.2.9 Orden de procesamiento de las imágenes

Tanto en el codificador como en el decodificador, el orden de procesamiento de las imágenes no coincide con el orden con el que dichas imágenes han sido adquiridas. En la terminología MPEG a este proceso se le conoce como ‘reordenación de tramas’.

El motivo de este aparente ‘descalabro’ de las imágenes se encuentra justificado por la propia filosofía del sistema de trabajo. Para comprenderlo, se van a razonar las situaciones sobre un ejemplo numérico concreto, ya que siempre es más fácil de entender los conceptos particularizados a unos números específicos.

Supóngase que se está trabajando con una relación entre imágenes I, P y B tales que se adaptan al ritmo 10/3, lo que quiere decir que cada diez imágenes se va a hacer una codificación tipo I y entre imágenes P se van a colocar dos imágenes tipo B. La situación en el orden de codificación es la que puede observarse sobre la Figura 7-26.

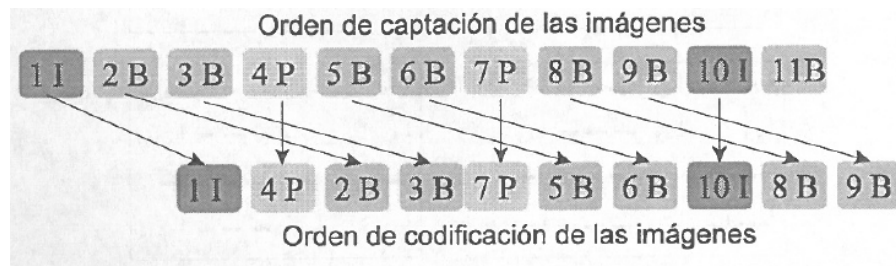


Figura 7-26. Alteración en el orden de codificación de las imágenes.

Es preferible alterar el orden de envío y hacerlo coincidente con el orden de codificación, porque de esta manera la cantidad de memoria necesaria en el decodificador es bastante menor.

7.3 Cuestionario

1-Los decodificadores MPEG-1 no podrán entender algunas variables recibidas y por tanto no procesarán esa señal.

- a) Verdadero.
- b) Falso.

2-Cuanto mayor sea el flujo binario:

- a) Menor será la calidad del vídeo.
- b) Mayor será la calidad obtenida en la imagen transportada.
- c) El régimen binario no influye en la calidad del vídeo.
- d) Ninguna de las opciones anteriores es cierta.

3-Los tipos de redundancia son:

- a) Temporal, espacial y de codificación.
- b) Temporal, espacial, de codificación y psicovisual.
- c) Temporal y psicovisual.
- d) Ninguna de las opciones anteriores

4-Un macrobloque es:

- a) Un conjunto de 16x16 píxeles.
- b) Un conjunto de 8x8 píxeles.
- c) Un conjunto de 16x16 bloques.
- d) Ninguna de las opciones anteriores es correcta.

5-Un bloque es:

- a) Un conjunto de 16x16 píxeles.
- b) Un conjunto de 8x8 píxeles.
- c) Un conjunto de 16x16 bloques.
- d) Ninguna de las opciones anteriores es correcta.

6-El formato de muestreo 4:4:4:

- a) Indica que en cada línea se desprecian la mitad de las muestras de crominancia, eligiendo una sí y otra no.
- b) Indica que sólo se desprecian las muestras de luminancia.
- c) Indica que no se ha despreciado ninguna de las muestras de crominancia y que por lo tanto se utilizan todas
- d) Ninguna de las opciones anteriores es cierta.

7-En la transformada discreta del coseno el valor de $\beta(v)$ cuando v es cero es: $(1/\sqrt{2})$

8-La codificación tipo inter:

- a) Usa imágenes anteriores y/o posteriores a la actual para codificar solamente la diferencia entre las mismas ahorrando una gran cantidad de bits.

- b) Usa sólo imágenes anteriores a la actual para codificar solamente la diferencia entre las mismas ahorrando una gran cantidad de bits.
- c) Usa sólo imágenes posteriores a la actual para codificar solamente la diferencia entre las mismas ahorrando una gran cantidad de bits.
- d) Ninguna de las opciones anteriores es cierta.

9-El proceso de codificación-decodificación es simétrico.

- a) Verdadero.
- b) Falso.

10-Un GOP integra imágenes:

- a) Integra imágenes sólo intra-codificadas.
- b) Integra 10 imágenes Intra y 10 Inter.
- c) Integra imágenes Inter e Intra.
- d) Ninguna de las opciones anteriores es correcta.

11-La estructura jerárquica de la señal de vídeo es por este orden:

- a) GOP, imagen, franja y macrobloque.
- b) GOP, imagen, franja, macrobloque y bloque.
- c) GOP, imagen y franja.
- d) Ninguna de las opciones anteriores es cierta.

12-El orden de procesamiento de las imágenes difiere al del orden lógico:

- a) Debido a que hay imágenes que necesitan ser procesadas en distinto orden como dictan las imágenes tipo B.
- b) No importa el orden de las imágenes cuando van a ser codificadas.
- c) Debido a que hay imágenes que necesitan ser procesadas en distinto orden como dictan las imágenes tipo Intra.
- d) Ninguna de las opciones anteriores es cierta.

13-Una distribución típica de imágenes es:

- a) 1/12
- b) 15/3
- c) 16/150
- d) 13/2

14-La Norma MPEG es útil solamente para vídeo.

- a) Verdadero.
- b) Falso.

15- Según la codificación *Run-Length* la cadena abbbbbbbccdddeeddd será:

- a) 1a7b2c4d2e3d.
- b) 1a7b3c4d2e3d.

c) 1a7b4c4d2e3d.

d) 1a7b5c4d2e3d.

8 SOPORTES DE ALMACENAMIENTO DE AUDIO Y VÍDEO

Para audio y vídeo se usan los mismos soportes de grabación que para cualquier otro tipo de dato, ya que bien en soporte magnético, sólido, óptico o magneto-óptico, el sonido y la imagen no son más que un cúmulo de datos (desde el punto de vista del almacenamiento).

Se pueden dividir en varios grupos de soportes atendiendo a diferentes características, como naturaleza tecnológica del soporte, velocidad de grabación y de transferencia, precio por byte, durabilidad, capacidad máxima, portabilidad, fiabilidad, etc... Cada uno dispone de una serie de ventajas e inconvenientes, que lo hacen mejor o peor en función de las necesidades de transferencia y almacenamiento.

En la actualidad existen multitud de estructuras y materiales con los que se fabrican soportes de almacenamiento. Según su construcción y los medios usados se pueden encontrar memorias:

Estado sólido (SSD)

Es un dispositivo de almacenamiento de datos que usa una memoria no volátil. En comparación con los discos duros tradicionales, las unidades de estado sólido son menos sensibles a los golpes, son prácticamente inaudibles y tienen un menor tiempo de acceso y de latencia. Las SSD hacen uso de la misma interfaz que los discos duros y, por lo tanto, son fácilmente intercambiables sin tener que recurrir a adaptadores o tarjetas de expansión para compatibilizarlos con el equipo. Este grupo incluye memorias internas, tarjetas compact flash, SD, *memory stick* o las tarjetas XD.

Ópticas

Las memorias en disco óptico almacenan información usando agujeros minúsculos grabados con un láser en la superficie de un disco circular. La información se lee iluminando la superficie con un diodo láser y observando la reflexión. Los discos ópticos son no volátiles y de acceso no secuencial. Los siguientes formatos son de uso común: CD, CD-ROM, DVD, CD-R, DVD-R, DVD+R, CD-RW, DVD-RW, DVD+RW, DVD-RAM, Blu-ray o HD DVD.

Magneto-ópticas

Las memorias de disco magneto óptico son un disco de memoria óptica donde la información se almacena en el estado magnético de una superficie ferromagnética. La información se lee ópticamente y se escribe combinando métodos magnéticos y ópticos. Las memorias de discos magneto ópticos son de tipo no volátiles, de acceso secuencial, de escritura lenta y lectura rápida.

Magnéticas

Las memorias magnéticas usan diferentes patrones de magnetización sobre una superficie cubierta con una capa magnetizada para almacenar información. Las memorias magnéticas son no volátiles. Se llega a la información usando uno o más cabezales de lectura/escritura. Como el cabezal de lectura/escritura sólo cubre una parte de la superficie, el almacenamiento magnético es de acceso secuencial y debe buscar, dar vueltas o las dos cosas. En computadoras modernas, la superficie magnética será de alguno de estos tipos: disco magnético, disquete, disco duro o cinta magnética (las antiguas cintas de vídeo o los magnetófonos).

Uno de los valores a tener en cuenta a la hora de escoger un sistema de almacenamiento es la capacidad. La unidad de medida son los Bytes, si bien los valores actuales alcanzan cifras de Mibibytes (MiB), Gibibytes (GiB) y Tebibytes (TiB). Cuanto mayor sea la capacidad, mayor será la cantidad de información almacenada, estando los distintos órdenes de magnitud directamente relacionados con las tecnologías de fabricación. Los sistemas magnéticos como los discos duros ya ofrecen valores de varios TiB lo cual los sitúa como firmes candidatos a almacenamiento masivo. Los mejores sistemas ópticos ofrecen valores de unos 50 GiB a un precio bastante competitivo, lo que permite crear sistemas de alta capacidad escalables a coste intermedio. Los modelos de

memoria Flash son los que ofrecen valores más bajos de entre 1 y 16 GiB, si bien existen sistemas que alcanzan hasta 256 GiB a precio alto.

La velocidad de funcionamiento se mide en b/s (bits por segundo o *bit per second*) y también es conocida como tasa de transferencia de datos. Se puede considerar por separado la velocidad de lectura y la de escritura. Los más rápidos son los sistemas Flash que pueden alcanzar velocidades de 712 Mb/s, frente a los 300 Mb/s de los HDD SATA II o los 54 Mb/s de los discos ópticos Blu-Ray.

El tamaño físico es un factor importante en los soportes, ya que el uso de muchos dispositivos implica la portabilidad del equipo completo, así como una cierta independencia eléctrica. Las tarjetas de memoria Flash, con un tamaño similar a una moneda, son usadas directamente en las cámaras y representan la opción más portátil y cómoda.

El precio viene definido por el coste de un determinado sistema dividido por la capacidad del mismo. Los HDD (Hard Disk Drive) pueden tener un precio por byte muy bajo, del orden de 0,00005 €/MiB, que los hace ideales para el almacenamiento masivo. Las tarjetas SD (Secure Digital) de máxima velocidad pueden disparar esta relación hasta los 0,006 €/MiB (casi 100 veces más). Los sistemas con precio por byte más bajos son utilizados para el almacenamiento final, mientras que los más caros están dedicados al almacenamiento intermedio.

La fiabilidad de un sistema de almacenamiento es la que define su buen funcionamiento, permitiendo la grabación de datos así como el acceso a los mismos en repetidas ocasiones con mínima tasa de error. Es recomendable utilizar soportes que garanticen la integridad de la información.

La reutilización del sustrato es la característica que permite distinguir entre sistemas de uno o más usos. Las memorias Flash son un soporte claramente reutilizable. Los discos ópticos, aunque hay modelos regrabables, son un ejemplo de soporte de un sólo uso. Los sistemas magnéticos también son soportes reutilizables.

En este capítulo se atenderá a la naturaleza tecnológica de cada soporte para hacer un estudio más en profundidad.

8.1 Memorias en estado sólido

Utilizan tecnologías de fabricación de tipo Flash RAM (Random-Access Memory), similares a las utilizadas en las memorias de ordenador y, en general, memorias basadas en chips de silicio. Actualmente son las más utilizadas para almacenar imagen y sonido capturado directamente desde el dispositivo, ya que son las más rápidas, resistentes y compactas. Dado que el precio por byte es elevado, su uso se limita al almacenamiento temporal, hasta su descarga en otro más adecuado para su procesado y archivado.

Entre las ventajas de este tipo de memorias se puede destacar su reducido tamaño, la elevada velocidad de transferencia con la que trabajan, el bajo consumo, su baja tasa de fallos, la resistencia que tienen a los golpes y a temperaturas extremas, desde -25 °C hasta 85 °C y que prácticamente no emiten ruido.

En contraposición tienen cierta limitación en la capacidad máxima de almacenamiento, un elevado precio por byte y un número limitado de escrituras y borrados, generalmente entre diez mil y un millón.

8.1.1 Memorias internas

La mayoría de los dispositivos electrónicos, y especialmente los capturadores digitales, cuentan con una pequeña cantidad de memoria integrada. Las cámaras que usan memoria interna fueron hace tiempo una opción bastante extendida, cuando los archivos ocupaban unos pocos MiB y permitían almacenar algunas decenas de imágenes. Actualmente es un sistema en desuso, salvo en algunos grabadores digitales de voz o cámaras de gama media-baja con orientación doméstica. Sus principales inconvenientes son la obligatoriedad de conectar directamente el dispositivo a un ordenador para descargar las grabaciones y la limitación de capacidad.

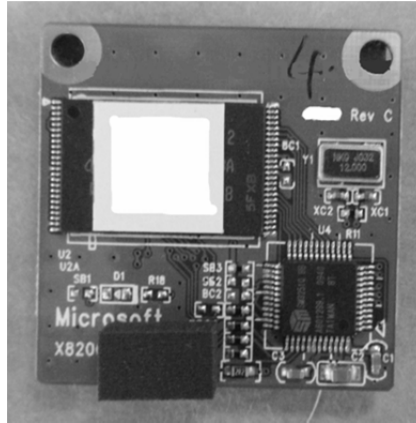


Figura 8-1. Memoria en estado sólido de Microsoft Xbox 360.

8.1.2 Tarjetas compact flash CF

Es el tipo de tarjeta de memoria que lleva más tiempo de los actualmente disponibles. Fue presentado en 1994 por la firma Sandisk Corporation como una evolución y miniaturización de las PCCARD (PCMCIA-RAM Tipos I a IV). Su reducido volumen y el hecho de no necesitar alimentación interna, fueron los factores determinantes para su aceptación y su masiva integración por parte de los fabricantes. Las ventajas de este tipo de dispositivos es la elevada resistencia física, su gran capacidad de almacenamiento pudiendo llegar a los 128 GiB, las altas velocidades de transferencia, que en modelos profesionales puede llegar a alcanzar incluso 100 MB/s permitiendo la grabación de vídeo Full HD 1080. Principalmente hay dos tipos de tarjetas CF, el Tipo I y el Tipo II ligeramente más grueso. Están disponibles en tres tasas de transferencia:

- CF original de 100 Mb/s.
- CF 2.0 (o CF+) de hasta 16 MB/s.
- CF 3.0 que alcanza los 66 MB/s e incluso más.

Son unos de los sistemas más “duros” ya que resisten golpes, vibraciones e incluso inclemencias meteorológicas mejor que otros tipos (-45 °C hasta 85 °C). Está entre las de mayor gasto energético dentro de los sistemas de memoria de estado sólido.

Para la conexión al ordenador utiliza la interfaz IDE/ATA y la USB en los modelos CF+.



Figura 8-2. Tarjeta compact flash de SanDisk.

Existe un formato especial desarrollado en 1999 por IBM llamado *Microdrive* que pese a utilizar el encapsulado CFII, utiliza tecnología de discos magnéticos, es decir micro discos duros. Como estos dispositivos encajan en cualquier ranura CFII, reciben más energía por término medio que la memoria flash, por lo que no pueden trabajar en algunos dispositivos con un nivel bajo de energía. Al tratarse de dispositivos mecánicos, son más sensibles a movimientos bruscos y cambios de temperatura que las memorias flash, aunque en la práctica son muy robustos.

8.1.3 Tarjeta SD (Secure Disk)

Desarrolladas por Panasonic en la década de los 90 como evolución del sistema MMC (Multimedia Card) al que se le integraba un sistema de cifrado ‘secure’ desarrollado por Toshiba. Son el sistema de memoria intercambiable más utilizado hoy en día, debido a su excelente relación de tamaño físico/capacidad, junto con su bajo consumo y su elevada velocidad de grabación. Al contrario que otras tarjetas de memoria, las SD también son ampliamente utilizadas por otros dispositivos electrónicos como teléfonos móviles, videoconsolas, ordenadores y *tablets*.



Figura 8-3. Tarjeta SD de SanDisk.

Las tarjetas de baja velocidad soportan tasas de transferencia de hasta 400 kb/s y modo de transferencia un-bit SD, mientras que las tarjetas de alta velocidad soportan tasas de transferencia de 0 a 100 Mb/s en el modo de cuatro-bits, y de 0 a 25 Mb/s en el modo un-bit.

Han sufrido varias evoluciones retro-compatibles que las han mejorado en capacidad y velocidad:

- Las versiones SDHC (High Capacity) alcanzan 128 GiB.
- Las más recientes SDXC (Xtended Capacity) elevan su capacidad hasta los 2 TiB.
- Otras evoluciones son las versiones de tamaño reducido MiniSD y MicroSD.

La velocidad mínima de transferencia viene definida por la “CLASE” de la tarjeta, de manera que una tarjeta asegura una transferencia de tantos Mb/s como el valor de su clase: clase 2 ofrece 2 Mb/s, clase 4 ofrece 4 Mb/s, clase 10 un mínimo de 10 Mb/s.

8.1.4 Tarjeta Memory Stick

Dentro de dicha familia se incluye la *Memory Stick Pro*, una versión posterior que permite una mayor capacidad de almacenamiento y velocidades de transferencia de archivos más altas. La *Memory Stick Pro Duo* es una serie nueva que permite mayor velocidad de lectura y transferencia y la *Memory Stick Pro-HG Duo*, una versión de menor tamaño que el *Memory Stick*. La *Memory Stick Micro* o M2 de tamaño similar a una microSD card y muy empleada en teléfonos móviles. En la actualidad diversas empresas han comercializado adaptadores de *Memory Stick Pro Duo* y *Memory Stick* que permiten emplear tarjetas de memoria microSD card en los dispositivos diseñados para una *Memory Stick*.

Su funcionamiento es similar al de las SD con un consumo bastante reducido y velocidades de transferencia similares, que pueden alcanzar hasta los 15 Mb/s lo que implica que garantizan una velocidad mínima mayor que las SD. Son bastante resistentes a golpes, vibraciones y temperaturas extremas.



Figura 8-4. Memory Stick ProDuo de Sony.

La *Memory Stick* incluye un amplio rango de formatos actuales, incluyendo dos factores de forma diferentes. La *Memory Stick* original era aproximadamente del tamaño y espesor de una goma de mascar y venía con

capacidades de 128 MiB hasta 2 GiB. Después Sony introdujo *Memory Stick PRO Duo*, de más capacidad que la *memory stick* normal, llegando a tener entre 4 y 32 GiB.

8.1.5 Tarjeta XD

Presentada por Olympus y Fujifilm en 2002, fue diseñada para albergar grandes archivos de imagen y vídeo, pudiendo alcanzar de forma teórica hasta 8 GiB desde su primera versión.



Figura 8-5. Tarjeta XD de Toshiba.

Necesita que el dispositivo (cámara,...) reconozca la capacidad y el protocolo de transmisión de datos, lo cual le aporta una mayor versatilidad de fabricación y un menor coste. Es un formato ideal para dispositivos compactos por su reducido tamaño, ya que es el formato más pequeño de todos los existentes (sin contar formatos reducidos como MiniSD, MicroSD, MS Micro y M2).

No ha logrado implantarse en el mercado y actualmente ni las propias marcas impulsoras lo incluyen en toda su gama. Tampoco ha evolucionado a sistemas de mayor velocidad o capacidad.

8.2 Soportes ópticos

La mayoría de sistemas de almacenamiento se basan en el registro de datos sobre la superficie lisa de un disco. Esto permite el acceso aleatorio a la información así como una mayor facilidad a la hora de generar copias exactas de un mismo soporte.

Debido a que la escritura y lectura de los datos se realiza de forma óptica no existe contacto entre la superficie del disco y el sensor óptico, lo cual minimiza el efecto de erosión o de ruptura de ambos elementos. Además este tipo de soportes pueden albergar en una misma superficie varias capas de información, como se aprecia en la Figura 8-6, e incluso tener disponibles para almacenamiento, ambas caras de un mismo disco. Sin embargo al poder almacenar tanta información los datos almacenados han de ir fuertemente protegidos frente a errores, lo cual produce un aumento considerable en los bits a tratar por los grabadores o reproductores ópticos.

Los discos ópticos son leídos por un rayo láser que hace las veces de cabeza lectora.

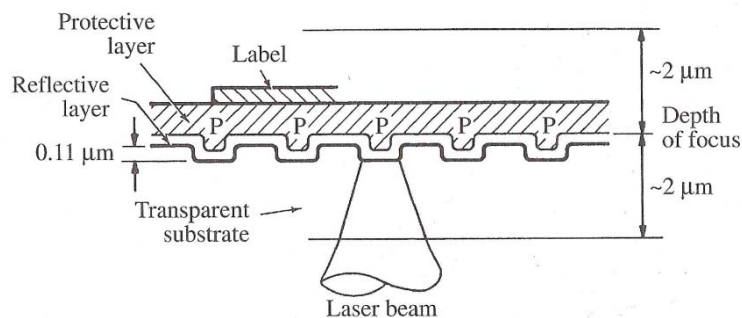


Figura 8-6. Ejemplo de capas del Disco Óptico Compacto (CD).

Cada capa está formada por varias superficies o capas de materiales:

1. Substrato: separa el exterior del disco de la siguiente capa. El láser atraviesa el sustrato sin apenas variar su trayectoria.

2. Capa de compuestos de grabación: las características de estos compuestos son modificadas durante la grabación pasando a contener la información grabada.
3. Capa reflectante: permite que el rayo emitido se refleje en el disco, atraviese el sustrato y sus variaciones sean leídas por el sensor situado cerca del emisor.
4. Capa de recubrimiento: protege las capas anteriores por la parte contraria al sustrato.

Los datos pueden ser almacenados o bien en una pista espiral o en círculos concéntricos, además pueden ser representados como un cambio de fase, de polarización o de intensidad de la luz reflejada del láser.

Todo material de grabación óptica ha de tener estabilidad a largo plazo (de hecho la vida útil de estos soportes ronda los 100 años), una alta absorción de la longitud de onda de grabación, baja energía de grabación, una alta SNR, una baja conductividad térmica y tiene costes bajos de manufacturación.

Los discos ópticos en general tienen un diámetro de entre 7.6 y 30 cm, siendo 12 cm el tamaño más común. Un disco típico tiene un grosor de 1.2 mm, mientras que el largo de pista, la distancia desde el centro de una pista hasta el centro de la siguiente, es en general de 1.6 μm , como muestra la Figura 8-7.

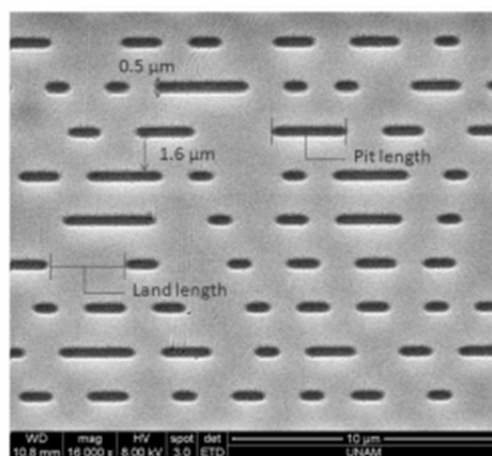


Figura 8-7. Cada línea horizontal son las pistas de un disco óptico, en este caso el de un CD.

Un disco óptico está diseñado para soportar uno de tres tipos de grabación: sólo lectura, por ejemplo CD y CD-ROM, grabable, posibilidad de escribir una sola vez, por ejemplo CD-R, o regrabable, por ejemplo CD-RW. Los discos grabables usualmente poseen una capa de grabación de tinte orgánico entre el sustrato y la capa reflexiva. Por otra parte, los discos regrabables contienen una capa de grabación de aleación compuesta de un material en cambio de estado, la mayoría de las veces aleación de plata, indio, antimonio y telurio.

El formato disco óptico cuenta con el respaldo de la ECMA, la organización que desde 1984 es responsable de avanzar en el desarrollo de los discos ópticos, al que pertenecen todos ellos: CD-ROM, DVD, UMD, Minidisc o *Blu-Ray*.

8.2.1 Disco compacto o CD (Compact Disk)

El sistema óptico fue desarrollado por Philips mientras que la lectura y codificación digital corrió a cargo de Sony, fue presentado en junio de 1980 a la industria, y se adhirieron al nuevo producto 40 compañías de todo el mundo mediante la obtención de las licencias correspondientes para la producción de reproductores y discos.

Un disco compacto contiene digitalmente codificados datos que son leídos por un haz láser. Debido a que la capa reflectante está incrustada en el disco, el polvo y las huellas que quedan en la superficie no afectan en la reproducción y los errores que se puedan producir se compensan con algoritmos correctores de errores. Pero el CD sobretodo representa un nuevo estándar de calidad sin precedentes para el consumo con una respuesta en frecuencia plana y baja distorsión en comparación con los discos de vinilo analógicos. Pero aun con todas las ventajas, sus sucesores superan todas sus características. Estos son el Super Audio CD, DVD y el *Blu-ray*.

Un CD almacena señal de audio estéreo comprimiendo dos palabras de 16 bits muestreadas a 44,1 kHz, lo cual resulta un total de 1,42 Mb/s de datos de audio más datos necesarios para la lectura. Todo junto, audio y datos hacen que el canal de salida del CD a la hora de leerlo sea de 4,32 Mb/s. En total un CD contiene una hora de música ya que soporta unos 15 Mb de canal de los cuales 6,3 Mb son audio y el resto sobrecarga (de los cuales el 33% son para la corrección de errores, el 7% de sincronización y el 4% de control y visualización).

Un disco compacto estándar mide 12 cm de diámetro y tiene un tiempo máximo de reproducción de 74 minutos y 33 segundos, aunque se pueden alcanzar tiempos algo mayores. Por ejemplo con una pista de 1,5 μ m a una velocidad lineal de lectura de 1,2 m/s se podrían alcanzar 82 minutos.

8.2.1.1 Formatos de CD

- **CD-ROM: *Compact Disc Read Only Memory*.** 600 - 700 MB. En este tipo de disco los datos están agrupados en sectores de 2352 bytes (24 bytes x 98 frames). Tiene 120 mm de diámetro, 1,2 mm grosor y un agujero central 15 mm. Los pits son 0-12 mm en profundidad y alrededor de 0,6 mm de ancho, las espirales están separadas 1,6 mm entre sí, dando una densidad de pista de 16000 pistas/pulgada. Puesto que el CD-ROM está basado en la estructura del CD-DA, son más lentos que los discos duros. Con el fin de encajar el máximo de música en el disco, el estándar se definió sobre la base de almacenar los datos a la misma densidad lineal en el exterior que en el centro, por lo que hay más datos en las pistas exteriores que en las interiores. Debido a lo anterior, la velocidad angular del disco decrece cuando la cabeza se mueve desde el centro a las pistas exteriores. Por lo que para permitir el acceso aleatorio necesita acelerar y decelerar el disco, siendo este el mayor obstáculo para incrementar la velocidad.
- **CD-i: *Compact Disc interactive*.** Permite almacenar audio, vídeo, gráficos, texto y datos. La estructura del disco CD-i:
- **Photo CD:** Tecnología de digitalización de la imagen para almacenar, manipular y visualizar imágenes fotográficas. El escaneado de imagen es con 8 bits de luminancia, 8 bits de crominancia y compresión 4:1. El Photo CD se puede leer en un reproductor CD-ROM XA, Reproductor CD-i y Reproductor Photo CD.
- **CD-DA: *Compact Disc Digital Audio*.** Aparece en 1982, mejorando la calidad de sonido (sin chasquido ni silbidos de analógicos) el acceso aleatorio a cualquier parte del disco y además se puede almacenar mucha información en poco espacio (12 cm). El tiempo de almacenamiento es de 74' 33'' aunque se pueden alcanzar tiempos algo mayores. Por ejemplo con una pista de 1,5 μ m a una velocidad lineal de lectura de 1,2 m/s se podrían alcanzar 82 minutos.

La cuantificación es de 16 bits uniforme, el ancho de banda de 5 a 20000 Hz y el muestreo a 44.1 kHz de 2 canales estéreo independientes. La tasa binaria efectiva es $44100 \times 16 \times 2 = 1,42$ Mb/s y la tasa binaria total 4,32 Mb/s. Lleva corrección de errores tipo CIRC (Cross Interleave Reed-Solomon Code) y codificación de canal EFM (Eight-to-Fourteen Modulation). En total un CD contiene una hora de música ya que soporta unos 15 Mb de canal de los cuales 6,3 Mb son audio y el resto sobrecarga (de los cuales el 33% son para la corrección de errores, el 7% de sincronización y el 4% de control y visualización).

- **CD-R: *Compact Disc Recordable*** (1988). Registro permanente, se puede leer de forma indefinida. No permite ni borrado ni escritura de nuevos datos. Se añade una capa de registro con una sustancia orgánica coloreada. La capas de este tipo de disco se especifican en la Figura 8-8.

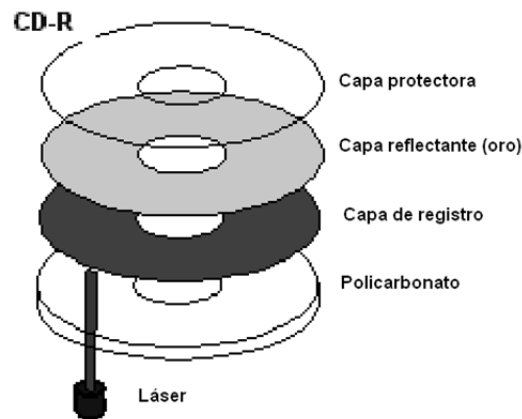


Figura 8-8. Capas de un CD-R.

Las pistas están en un surco en espiral preformado. El objetivo ha de guiar el haz láser en escritura y lectura. Permite al grabador controlar la velocidad de rotación del disco para obtener velocidad lineal constante (1.2 m/s).

El mecanismo de grabación es por calentamiento. El láser aumenta su potencia y calienta la capa de tinte orgánico. Éste atraviesa el policarbonato y calienta el tinte a 250 °C. El calentamiento provoca la descomposición química del punto disminuyendo la reflectividad del 75% al 25%, lo cual provoca un “pit simulado”.

- **CD-RW: Compact Disc Re-Writable** (1996). En este tipo de disco los datos pueden ser grabados, leídos y borrados muchas veces. Tiene dos capas más que el CD-R que son de dieléctrico, como se puede observar en la Figura 8-9.

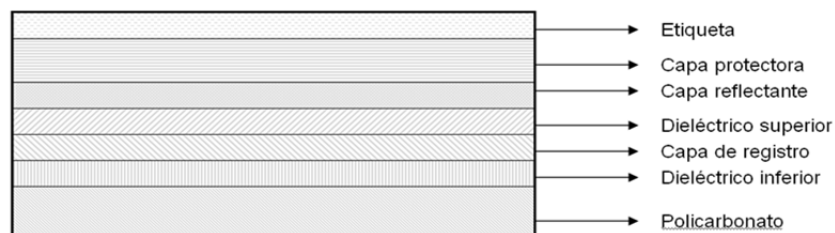


Figura 8-9. Capas del CD-RW.

El mecanismo de grabación está basado en tecnología de cambio de fase. Los discos regrabables poseen una capa de un compuesto químico formado de plata, antimonio, telurio e indio. Este compuesto químico tiene la particularidad de volverse líquido a temperaturas cercanas a 700 °C, y luego al enfriarse adquiere las propiedades de un sólido amorfo si estructura definida y **opaco**.

Si el compuesto se calienta hasta una temperatura inferior (200 °C aproximadamente), alcanza su temperatura de cristalización, donde se presenta como una estructura sólida definida y **traslúcida**.

El compuesto en estado amorfo es opaco y absorbe la mayor parte de la luz que incide en él, mientras se encuentra en estado sólido es translúcido. Esto, junto con la capa metálica reflectante que se encuentra sobre esta capa de compuesto hace que se comporte de forma similar a un CD-R.

8.2.2 DVD

El formato DVD (*Digital Versatile Disc*) es el sucesor directo del CD y mejora respecto a este en capacidad de almacenamiento y permite una lectura y una escritura mucho más rápidas. Esta mejora tecnológica proporciona por ejemplo, al mundo del cine, la música y los programas informáticos abrir sus horizontes más allá de los formatos antiguos.

A comienzo de los años 1990, dos estándares de almacenamiento óptico de alta densidad estaban desarrollándose: uno era el multimedia compact disc (MMCD), apoyado por Philips y Sony; el otro era el Super Density (SD), apoyado por Toshiba, Time Warner, Matsushita Electric, Hitachi, Mitsubishi Electric, Pioneer, Thomson y JVC.

Philips y Sony acordaron con Toshiba adoptar el SD, pero con una modificación: la adopción del EFM Plus de Philips, que a pesar de ser un 6% menos eficiente que el sistema de codificación de Toshiba (de ahí que la capacidad sea de 4,7 GB en lugar de los 5 GB del SD original), cuenta con la gran ventaja de que EFM Plus posee gran resistencia a los daños físicos en el disco, como arañazos o huellas. El resultado fue la creación del Consorcio del DVD, fundada por las compañías anteriores, y la especificación de la versión 1.5 del DVD, anunciada en 1995 y finalizada en septiembre de 1996.

Aunque está basado en la tecnología del CD, el DVD emplea nuevas especificaciones y nuevos formatos de archivos. El CD originalmente fue diseñado exclusivamente para almacenamiento de audio y después avanzó ligeramente hacia otros formatos, pero el DVD fue diseñado desde un principio como un formato de grabación para cualquier formato.

En todos los aspectos el DVD supera al CD: capacidad, rapidez, es más robusto, útil para cualquier formato, mayor régimen binario y ancho de banda, etc. Un DVD de sólo una capa tiene una capacidad de siete veces la de un CD. Esto es posible por las mejoras en el diseño óptico y sistemas más sofisticados de decodificación electrónica.

8.2.2.1 Diseño y características del DVD

La parte 1 de las especificaciones del DVD definen las características físicas de éste. Los formatos derivados del DVD, DVD-ROM, DVD-Video y DVD-Audio tienen exactamente las mismas características.

Donde el CD usa una capa simple de policarbonato el DVD emplea 2 substratos unidos con las capas de datos ubicados cerca de la capa interna. Todos los datos están grabados en una pista espiral. La capa más cercana a la superficie de lectura se denomina Layer 0, que se lee primero, y la capa más alejada es Layer 1. En la

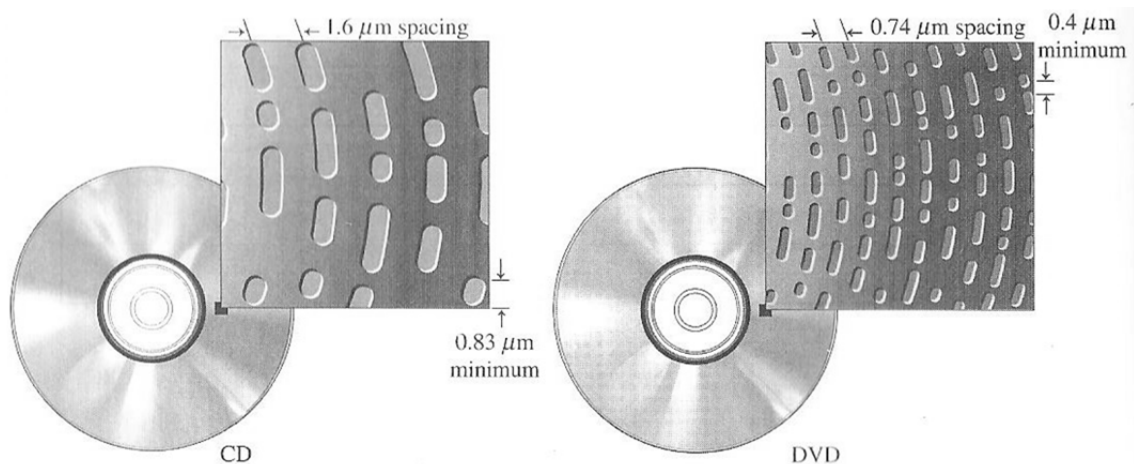


Figura 8-10. Comparación entre el CD y el DVD. Se puede apreciar que el DVD contiene hasta cuatro veces más pits que un CD. Esto sumado a las mejoras de los algoritmos de codificación y corrección de errores hace que la capacidad de un DVD llegue hasta ser siete veces mayor que la de un CD.

Cuando el láser hace la lectura empieza por la parte interna de la Layer 0 hasta la parte externa del CD, donde encuentra un área intermedia, para, y hace un reenfoque para comenzar a leer la Layer 1 desde la parte externa del CD hasta la interna. Esto se denomina organización inversa de encaminamiento de pista (*Opposite Track Path-OTP*). Una alternativa sería escribir la Layer 1 de forma independiente desde una parte interna a la externa del disco, esto es un encaminamiento paralelo de pista (*Parallel Track Path*).

La protección de la capa de datos está ubicada más al interior de lo que lo estaba en el CD. Dado que las capas son más finas, el DVD es menos sensible a los errores de salto de pista pero si es más fácil que aparezcan errores por la contaminación de la superficie de datos que en el CD, lo cual se compensa con unos mejores algoritmos de corrección de errores.

Los DVD se pueden clasificar según su contenido:

- **DVD-Vídeo:** películas (vídeo y audio). Aparece en el mercado en 1996. Sustituye al VHS. Su antecedente son el Vídeo-CD y el LaserDisc y es el predecesor del Blu-ray. Almacena hasta 133 minutos o más de alta calidad MPEG-2 con sonido *surround* multicanal por cada cara del disco DVD. Tiene diferentes formatos de visualización de pantalla, 8 idiomas diferentes, subtítulos en más de 32 idiomas, menús, acceso aleatorio, navegación y protección anticopia digital y analógica. Es soportado por los formatos DVD-5, DVD-9, DVD-10 y DVD-18. Permite acomodar hasta 8 canales independientes de audio, pudiendo ser: de 1 a 8 canales en PCM lineal, de 1 a 6 canales de la configuración envolvente 5.1 o de Dolby Digital y de 1 a 8 canales codificados en MPEG-2 AAC.
- **DVD-Audio:** Conocido como DVD-A. Por ejemplo: 24 bits por muestra, una velocidad de muestreo de 48 000 Hz y un rango dinámico de 144 dB. Es un formato de audio de alta fidelidad en DVD. El DVD-A está dividido en tres grandes zonas: Audio, Vídeo (videoclips) y Datos (textos, como por ejemplo letras de las canciones, biografías de los artistas, imágenes...). A estas zonas se accede mediante un menú, por lo que hay que conectar un monitor de televisión. Utiliza la modulación PCM para codificar el audio (de forma digital). Admite 6 frecuencias de muestreo: 44,1 kHz (y múltiplos 88,2 kHz y 176,4 kHz), 48 kHz (y múltiplos 96 kHz y 192 kHz). La respuesta en frecuencia del DVD-A depende de la frecuencia de muestreo. Los reproductores incorporan un filtro pasa bajo que atenúa progresivamente la señal a partir de los 20 kHz. El rango dinámico teórico que puede alcanzar el DVD-A con cuantificación de 24 bits es de unos 146,2 dB pero en la práctica es muy difícil que un conversor exceda una SNR de 120 dB.
- **DVD-Data:** todo tipo de datos.

8.2.2.2 Formatos de DVD

Según su capacidad de grabado:

- **DVD-ROM:** solo lectura, manufacturado con prensa. El DVD-ROM no es más que un gigantesco almacén de bits, preparado para recibir cualquier información que sea estructurada de acuerdo al estándar UDF (Universal Disk Format). Básicamente, todos los discos DVD son discos DVD-ROM. Los lectores DVD-ROM se conectan a un ordenador personal, funcionando como un tradicional lector de CD-ROM.
- **DVD-R y DVD+R:** grabable una sola vez. La diferencia entre los tipos +R y -R radica en la forma de grabación y de codificación de la información. En los +R los agujeros son 1 lógicos mientras que en los -R los agujeros son 0 lógicos. Permiten la posibilidad de grabar datos en el disco de forma permanente y en una sola ocasión. Los discos DVD-R contienen una zona destinada a la calibración del láser, denominada Área de Calibración de Potencia. El resto del disco comprende el Área de Información, estando dividida en tres zonas de datos.
- **DVD-RW y DVD+RW:** regrabable. Permite la reescritura de los datos, siendo sus especificaciones una extensión de las del DVD-R. El sistema utiliza un mecanismo de grabación por cambio de fase y una estructura de disco multicapa. Puede almacenar 4,7 GB por capa. No está diseñado como medio de almacenamiento y distribución de datos de tipo genérico, ya que es menos robusto en la protección frente a errores, y además cuenta con un número de ciclos de reescritura menor que otros sistemas de almacenamiento.
- **DVD-RAM:** regrabable de acceso aleatorio. Lleva a cabo una comprobación de la integridad de los datos siempre activa tras completar la escritura. Se diferencia del DVD-RW y del DVD+RW en que no hace falta borrar todo el disco para recuperar el espacio de los contenidos que deseamos borrar y en que se puede grabar directamente en él como si fuera un disquete, sin necesidad de programas de grabación de DVD, ni de programas controladores intermedios (en el caso de grabadores DVD-RAM para computadoras).
- **DVD+R DL:** grabable una sola vez de doble capa.

Según su número de capas o caras:

- **DVD-5:** una cara, capa simple, 4,7 GB o 4,38 GiB. Discos DVD±R/RW.
- **DVD-9:** una cara, capa doble, 8,5 GB o 7,92 GiB. Discos DVD+R DL. La grabación de doble capa permite a los discos DVD-R y los DVD+RW almacenar significativamente más datos, hasta 8,5 GB por disco, comparado con los 4,7 GB que permiten los discos de una capa. Un disco de doble capa difiere de un DVD convencional en que emplea una segunda capa física ubicada en el interior del disco. Una unidad lectora con capacidad de doble capa accede a la segunda capa proyectando el láser a través de la primera capa semitransparente. El mecanismo de cambio de capa en algunos DVD puede conllevar una pausa de hasta un par de segundos. Los discos grabables soportan esta tecnología manteniendo compatibilidad con algunos reproductores de DVD y unidades DVD-ROM. Muchos grabadores de DVD soportan la tecnología de doble capa, y su precio es comparable con las unidades de una capa, aunque el medio continúa siendo considerablemente más caro.
- **DVD-10:** dos caras, capa simple en ambas, 9,4 GB o 8,75 GiB. Discos DVD±R/RW.
- **DVD-14:** dos caras, capa doble en una, capa simple en la otra, 13,3 GB o 12,3 GiB. Raramente utilizado.
- **DVD-18:** dos caras, capa doble en ambas, 17,1 GB o 15,9 GiB. Discos DVD+R.

También existen DVD de 8 cm (no confundir con miniDVD, que son CD que contienen información de tipo DVD vídeo) que tienen una capacidad de 1,5 GB.

8.2.2.3 Codificación de datos

En la estructura de datos del DVD de sólo lectura los datos son almacenados en archivos lo que incrementa la densidad de datos, que son ubicados sobre el disco en sectores físicos que corren de manera continua sin dejar un hueco vacío.

El DVD emplea la codificación *Reed-Solomon Product Code* (RS-PC) para la corrección de errores. Este tipo de código usa una combinación de dos códigos *Reed-Solomon* (llamados C1 y C2) como un código producto. Se diferencia del CIRC usado en el CD, es más similar al del código usado en el DAT.

Los datos en el DVD son grabados con modulación EFMPPlus. Es un código 8/16 RLL (Run length Limited) que escribe cada byte de datos de 8 bits como un byte de canal modulado de 16 bits. El tipo de modulación de datos usado por el DVD incrementa un 6% la capacidad de almacenamiento de usuario comparándolo con el CD.

8.2.2.4 Proceso de lectura

Aunque la compatibilidad con los CD no es una obligación en los lectores de DVD, lo cierto es que prácticamente todos los reproductores de DVD pueden leer un CD. Para conseguir esto, el láser debe ser capaz de enfocar la capa de datos del CD (a 1,2 mm de la superficie de lectura) así como la capa de datos del DVD (a 0,6 mm de la superficie de lectura). Algunos lectores usan un solo objetivo con las longitudes de onda necesarias para el CD y el DVD, 780 nm y 635/650 nm respectivamente.

8.2.3 Blu-ray

El sistema de disco óptico Blu-ray fue desarrollado como el sucesor del formato DVD-Vídeo y como un competidor del fallido HD-DVD. El disco Blu-ray está orientado al consumo y es usado de manera generalizada en la reproducción de películas en alta definición. Su calidad es mayor que la del DVD de definición estándar y satisface los estándares DTV de emisión en alta definición. Además se puede grabar y reproducir vídeo estereoscópico (3D), puede ser usado en vídeo juegos y medios regrabables. En cuanto al audio, es capaz de reproducir audio multicanal de alta calidad de larga duración. El Blu-ray emplea protección de contenido más robusta que la usada en el DVD.

El grupo de compañías que desarrolló el disco Blu-ray, fue encabezado por Sony y fue fundado en 2002, aunque se renombró en 2004 como *Blu-ray Disc Association*. El nombre del disco deriva de la palabra azul en inglés (*blue*) pero al no poder establecerse nombres genéricos como marcas comerciales se optó por usar *Blu* en su lugar.

Al igual que el CD y el DVD, el disco *Blu-ray* está disponible en formatos pregrabados, grabables y regrabables, BD-ROM, BD-R y BD-RE respectivamente. Los tres tipos tienen la misma capacidad de almacenamiento y pueden soportar capas simples o capas dobles. Las capas son independientes y se leen desde la misma cara.

8.2.3.1 Capacidad del disco

Un disco Blu-ray de una sola capa BD-ROM, conocido como BD-25 puede almacenar alrededor de 25 GB, o lo que es lo mismo, es capaz de soportar hasta 2 horas de vídeo en alta definición. El BD-50 puede aguantar hasta 50 GB de datos. Existen diversos tamaños y capacidades para el Blu-ray que se especifican en la Tabla 8-1.

Disc Type	Diameter (cm)	Disc Structure	Capacity (billions of bytes) (10^9) ¹	Capacity (gigabytes) (2^{30}) ¹	Typical Video Playback Time (hours) ²	2-Channel Audio-Only Playback Time (hours) ³	5.1-Channel Audio-Only Playback Time (hours) ⁴
BD-8	8	1 layer	7.791	7.256	0.7	3.8	27.1
BD-16	8	2 layers	15.582	14.512	1.4	7.5	54.1
BD-25	12	1 layer	25.025	23.306	2.3	12.1	86.9
BD-27	12	1 layer	27.020	25.164	2.5	13.0	93.8
BD-50	12	2 layers	50.050	46.613	4.6	24.1	173.8
BD-54	12	2 layers	54.040	50.329	5.0	26.1	187.6

Tabla 8-1. Capacidades de almacenamiento y típicas duraciones de diferentes tipos de discos BD.

8.2.3.2 Diseño del disco

Al igual que el CD y el DVD, el disco Blu-ray almacena los datos binarios usando pits grabados en un sustrato. Durante la reproducción los pits provocan un cambio en la intensidad de luz reflejada en la cara de lectura y esos cambios, 980.000 por segundo, son decodificados para reproducir los datos que contiene el disco.

Tanto la capa simple como la capa doble tienen un sustrato con una anchura nominal de 1,1 mm. En discos de capa simple los datos se cubren con una capa reflectante cubierta a su vez por otra capa de 0,1 mm de anchura. En los discos de doble capa, éstas se separan por alrededor de 0,025 mm de material transparente. La capa más interna, L0, se cubre de una capa totalmente reflectante y la capa más externa, L1, se cubre con una capa semi-reflectante.

En el disco de Blu-ray las características del sustrato no son críticas ya que el camino óptico está grabado sobre la capa de cubrimiento, que proporciona resistencia a las ralladuras lo que hace menos necesario el uso de una carcasa. Los discos CD, DVD y Blu-ray tienen todos, una pista en espiral y se leen desde la parte más interna a la más externa del disco.

8.2.3.3 Lectura

Como ya se ha dicho la capa de cubrimiento sobre la capa de datos es más fina comparada con las del CD o el DVD, y la longitud de onda del láser mucho más corta. Por estas y otras razones no es posible usar en los discos Blu-ray el mismo sistema óptico de lectura usados en esos discos. Sin embargo usando múltiples elementos, algunos lectores son capaces de leer los tres formatos.

Códec de audio

En la Tabla 8-2 se especifican los Códec de audio que soporta el Blu-Ray.

Audio Codec	Mandatory	Compression	Sampling Frequency (kHz)	Word Length (bits)	Maximum Bit Rate (Mbps)	Maximum Number of Channels
PCM	Yes	Lossless	48, 96, 192 ¹	16,20,24	27.648	8 (48 kHz and 96 kHz) 6 (192 kHz)
Dolby Digital	Yes	Lossy	48	16–24	0.640	5.1
Dolby Digital Plus	No	Lossy	48	16–24	1.7	7.1 ²
Dolby TrueHD	No ³	Lossless	48, 96, 192 ¹	16–24	18.64	8 (48 kHz and 96 kHz) 6 (192 kHz)
DTS	Yes	Lossy	48	16,20,24	1.509	5.1
DTS High Resolution	No ⁴	Lossy	48, 96	16–24	6	7.1
DTS-HD Master Audio	No ⁴	Lossless	48, 96, 192 ¹	16–24	24.5	8 (48 kHz and 96 kHz) 6 (192 kHz)

Tabla 8-2. Códec soportados por el Blu-ray.

Códec de vídeo

En la Tabla 8-3 se especifican los Códec de vídeo que soporta el sistema Blu-ray.

Video codecs	MPEG-4 AVC: HP@4.1/4.0 and MP@4.1/4.0/3.2/3.1/3.0, SMPTE VC-1: AP@L3 and AP@L2, MPEG-2: MP@ML and MP@HL/H1440L
Picture (HD)	1920 × 1080 × 59.94i, 50i (16:9) 1920 × 1080 × 24p, 23.976p (16:9) 1440 × 1080 × 59.94i, 50i (16:9) MPEG-4 AVC/SMPTE VC-1 only 1440 × 1080 × 24p, 23.976p (16:9) MPEG-4 AVC/SMPTE VC-1 only 1280 × 720 × 59.94i, 50i (16:9) 1280 × 720 × 24p, 23.976p (16:9)
Picture (standard definition)	720 × 576 × 50i (4:3/16:9) 720 × 480 × 59.94i (4:3/16:9)
Display aspect ratio	4:3 and 16:9
Frame size	24/23.976 fps, 29.97/59.94 fps, 25/50 fps (PAL regions)
Audio codecs	PCM, Dolby Digital, Dolby Digital Plus, ¹ Dolby TrueHD, ¹ DTS, DTS-HD High Resolution, ¹ DTS-HD Master Audio ¹
Audio channel configuration	1.0 to 7.1 channels
Audio sampling frequency	48, 96, 192 kHz
Subtitles and graphics	8-bit Interactive Graphics stream for menus, 8-bit Presentation Graphics stream for subtitles, HDMV Text Subtitle stream, 32-bit RGBA graphics for BD-J menus
Stream structure	MPEG-2 System Transport Stream
Maximum transport stream bit rate	48 Mbps
Maximum video stream bit rate	40 Mbps
Maximum number of video streams	9 primary/32 secondary
Maximum number of audio streams	32 primary/32 secondary
Maximum number of graphics streams	32
Maximum number of text streams	255

Tabla 8-3. Especificaciones para los códecs de vídeo soportados por el sistema de Blu-ray.

8.2.3.4 Modulación y corrección de errores

Los datos en un disco Blu-ray, exceptuando la palabra de bits de sincronización, se graban usando 1-7PP (*Parity preserve/Prohibit repeated minimum transition run length* modulation).

El algoritmo de corrección de errores usado en el disco Blu-ray es muy eficiente, por ejemplo los datos de corrección de errores consumen menos del 17% de la señal de datos, antes de la modulación la capacidad de corrección ya es considerable. En los discos Blu-ray, la capa de cubrimiento sobre la capa de datos es relativamente fina, lo que hace que la suciedad y las partículas de polvo afecten más en la lectura del disco. Esto hace que se precise de un sistema muy robusto de corrección de errores. Los discos Blu-ray combinan un profundo intervalo con un código de larga distancia (*Long Distance Code* - LCD), código Reed-Solomon con un tamaño de 64 kbytes con un indicador de ráfaga (*Burst Indicator Subcode* BIS).

8.3 Soportes magneto-ópticos

8.3.1 MiniDisc

El MiniDisc, miniDisk o MD, es un disco magneto-óptico digital desarrollado en los 90 por Sony, de menor tamaño que los CD convencionales y mayor capacidad en comparación. Actualmente ya se ha cesado en su distribución dada su baja demanda.

Es un disco óptico de pequeñas dimensiones (7 cm x 6,75 cm x 0,5 cm) y regrabable, de almacenamiento magneto-óptico diseñado inicialmente para contener hasta 80 minutos de audio digitalizado.

El formato que se usa para la música está basado en la compresión ATRAC/ATRAC3, usa DRM (*Digital Rights Management*), diferentes regímenes binarios, y un muestreo directo a partir de una señal digital o analógica.

Los discos MiniDisc son más pequeños que los CD, tienen un diámetro de 64 mm, pero su velocidad de transferencia de datos es menor: 292 kb/s, frente a los 1,4 Mb/s que requiere el CD. Entrega una resolución de 16 bits, utilizando para ello la frecuencia de muestreo estándar 44,1 kHz.



Figura 8-11. MiniDisc de Sony.

8.3.1.1 Proceso de grabación

El Minidisc utiliza un sistema de grabación digital de sonido. La grabación magneto-óptica es un sistema combinado que graba de forma magnética, pero reproduce de forma óptica.

Los datos se graban en el disco mediante lo que se conoce como recubrimiento de cambio de fase.

La superficie del disco, mientras se encuentra bajo la influencia de un campo magnético, es calentada por el láser que se utiliza durante la lectura del disco aumentando su potencia durante la grabación, el calor generado hace que la superficie del disco alcance una temperatura crítica conocida como el punto de Curie (cerca de 180 °C). Cuando la aleación metálica que conforma el disco alcanza esta temperatura, modifica su estado de cristalización. Aprovechando este cambio en el estado de cristalización, el flujo magnético alinea los cristales en direcciones opuestas.

Al enfriar el disco rápidamente, el magnetismo inducido permanece. Esto se debe a que al enfriar rápidamente, se refleja menos luz y el material ya no recrystaliza adecuadamente, por lo que no vuelve a su estado original. Así, la información binaria ha quedado almacenada permanentemente. Los fabricantes aseguran que es posible almacenar los datos durante 30 años sin distorsiones o pérdidas.

Como en un disco compacto, el Minidisc almacena la música en pistas.

8.4 Soportes Magnéticos

La grabación magnética consiste en la grabación de datos en una banda magnética que crea un flujo del campo magnético al circula una corriente por el bobinado de la cabeza grabadora. Este flujo magnetiza la cinta en movimiento. Los sistemas magnéticos se dividen en Disco Duro y Cinta Magnética.

8.4.1 Cinta magnética

La grabación magnética ha sido el principal sistema de almacenamiento de la señal analógica de audio durante los últimos 50 años. Si a la facilidad de leer, escribir y borrar una cinta magnética le añadimos que durante los últimos 50 años la densidad de almacenamiento se ha incrementado más de 750000 veces, es lógico que se haya convertido en un medio de almacenamiento de características únicas. El avance que tuvo la industria informática hizo que las cintas no sólo almacenasen señal analógica, sino que soportan perfectamente la señal digital, de hecho las cintas no son un problema en cuanto a soportar dicha señal, sino que condicionan más la cantidad tan alta de datos que tienen que almacenar.



Figura 8-12. Hay diferentes sistemas de formatos magnéticos como el VHS, el cassette o el disquete.

Fundamentos de grabación magnética

La teoría magnética permite establecer una ecuación aproximada para la frecuencia máxima que se puede grabar y reproducir en cualquier soporte magnético que es la ecuación (8-1):

$$f_{max} = \frac{V_{cp}}{2e} \text{ Hz. (8-1)}$$

Donde V_{cp} es la velocidad relativa entre la cabeza magnética y la pista magnética y e es el ancho del elemento sensor magnético (esto es, el 'entrehierro' en caso de cabezas magnéticas de inducción o el ancho en el caso de cabezas magnetoresistivas).

Dado que la frecuencia mínima es muy cercana a unos pocos Hz, se puede afirmar que la frecuencia máxima es igual al ancho de banda.

Siguiendo los conceptos básicos de transmisión digital, el límite de Nyquist para una señal digital binaria en banda base viene dado por la expresión (8-2):

$$B \geq \frac{R_b}{2} \text{ Hz. (8-2)}$$

Donde B es el ancho de banda y R_b el régimen binario. Estas dos ecuaciones combinadas nos pueden dar la velocidad necesaria de lectura escritura de una cinta magnética.

Hay que tener en cuenta otro aspecto sobre las ecuaciones anteriores es la duración y el mantenimiento de las cabezas lectoras. Los soportes de grabación que supongan desgaste de las cabezas por el rozamiento de éstas con la cinta deben tener en consideración que el efecto del desgaste es aumentar el entrehierro e . Así, el desgaste disminuye el ancho de banda disponible en el soporte respecto al que tenía al principio de su uso. Cuando el ancho de banda disponible quede por debajo del límite de Nyquist, la grabación/reproducción ya no será posible. Esto que se acaba de decir hace que lleguemos a dos conclusiones sobre el desgaste en las cabezas:

- Las cabezas deben cambiarse cada cierto número de horas por otras nuevas.
- El entrehierro inicial no debe ser muy estrecho, para que el desgaste proporcional no sea muy alto.

8.4.1.1 Métodos de grabación

Grabación Transversal. Para aprovechar la superficie grabable de la cinta magnética se diseñó un tambor giratorio con cuatro cabezas de grabación para almacenar vídeo, el cual requiere mayor cantidad información

almacenada a lo ancho de la cinta, mientras la cinta estaba en constante movimiento cada pista está ligeramente inclinada, muy pocos equipos utilizaron este método de grabación como las grabadoras Mark I y Mark II de Ampex nombraron esta tecnología de grabación como "Quádruplex".

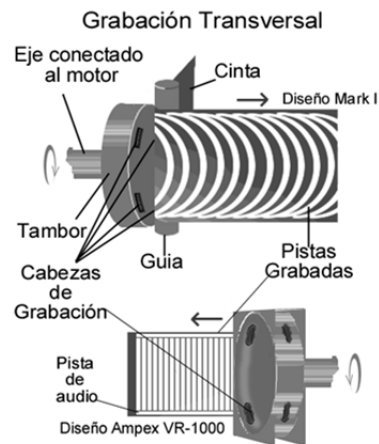


Figura 8-13. Esquema de la grabación transversal de la cinta magnética.

Grabación Lineal. La grabación de señales analógicas y digitales que se registran en la cinta magnética se realiza en pistas horizontales. Hubo variantes como la grabación en multipistas que comprenden múltiples trazas horizontales grabadas con diferentes cabezas que podían ser reproducidas al mismo tiempo, o el Lineal serpentine que al momento de llegar al final de la cinta la cabeza grabadora traza en dirección opuesta una segunda pista paralela lo que permitió una mayor capacidad de almacenamiento.

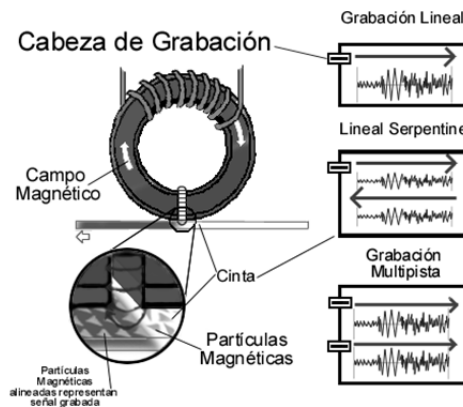


Figura 8-14. Esquema de grabación lineal en cinta magnética.

Grabación Helicoidal. Para perfeccionar el modo de grabación transversal se redujeron de cuatro cabezas a dos que rotan en diferentes ángulos hacia la dirección en la que la cinta se transporta, para lograr una secuencia de grabación continua y mayor superficie de cinta aprovechada, se pueden apreciar las trazas de la cinta grabadas diagonalmente a lo largo de esta.

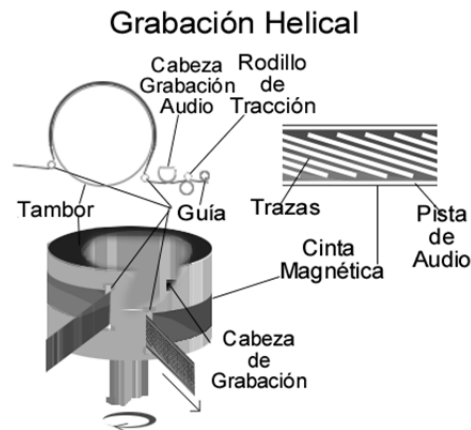


Figura 8-15. Esquema de grabación helicoidal en cinta magnética.

8.4.1.2 Especificaciones

El rango de frecuencias en el caso del magnetófono analógico es de 0 a 20 kHz, ideal para audio. Sin embargo, el registro digital requiere un ancho de banda mucho más elevado, de alrededor de 50 veces al de un magnetófono analógico ya que se pueden grabar 1 Mb/s, y aunque se pueda grabar una forma de onda distorsionada para elevar la densidad de almacenamiento de datos, el problema de la grabación de transiciones muy juntas no se llega salvar.

Una cinta magnética está compuesta por un soporte plástico, generalmente poliéster, revestido de una fina capa de material magnético. Estos materiales están compuestos por partículas de forma acicular de un tamaño de $0,3 \times 0,03 \mu\text{m}$ las de hierro gamma y $0,5 \times 0,04 \mu\text{m}$ las de cromo. En el proceso de fabricación de la cinta, las partículas son físicamente alineadas en la dirección de avance de la cinta, con lo que se incrementa la densidad de empaquetamiento. Cada partícula se comporta como un pequeño imán permanente, con la capacidad de producir un pequeño campo magnético; y la posibilidad de cambiar su orientación mediante un campo magnético externo. Antes de efectuar grabación alguna en la cinta (cinta virgen), los campos magnéticos de las partículas no aportan una orientación neta predeterminada. Para almacenar la información, un campo magnético externo orienta las partículas magnéticas, de forma que sus campos magnéticos se alinean con el campo externo aplicado. En la grabación analógica, la alineación neta relativa de los campos magnéticos de las partículas representa de forma continua la magnitud (amplitud) de la señal grabada; de forma que puede ser almacenado un cambio continuamente variable. , como se aprecia en la Figura 8-16B.

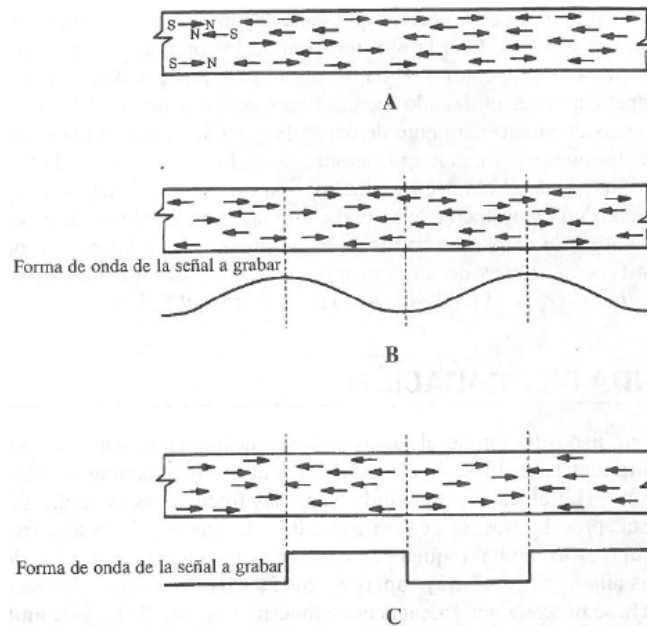


Figura 8-16. En el almacenamiento en cinta magnética, la polaridad de los campos magnéticos de las partículas de óxido se orienta de forma relativa a la amplitud y la frecuencia de la señal a grabar. A: Antes de la grabación, la polaridad de los campos magnéticos aparece orientada sin ninguna magnetización neta. B: El registro analógico provoca variaciones continuas. C: En la grabación digital, los datos están representados mediante saturación en cinta.

La grabación por saturación se emplea en el almacenamiento de datos binarios. La intensidad del campo magnético se incrementa para que todas las partículas de una zona sean orientadas. Cuando se aplica una forma de onda bipolar, la cinta saturada adquiere dos estados de igual magnitud pero de polaridad opuesta. La señal de escritura es una corriente que cambia de polaridad en transiciones de cadenas de bits. Las señales de la cabeza de escritura (cabeza de grabación) provocan que regiones enteras de la cinta sean orientadas de forma positiva o negativa, tal y como muestra la Figura 8-16-C. Estas transiciones en la dirección del campo magnético representan transiciones de 0 a 1 en los valores binarios de señal. A diferencia de la grabación analógica, en este caso no se necesita ninguna corriente de polarización que linealice el proceso, ya que aquí sólo importa la saturación del material.

8.4.2 Discos Duros

El Disco Duro es un dispositivo de almacenamiento de datos no volátil que emplea un sistema de grabación magnética para almacenar datos digitales. Se compone de uno o más platos o discos rígidos, unidos por un mismo eje que gira a gran velocidad dentro de una caja metálica sellada. Sobre cada plato, y en cada una de sus caras, se sitúa un cabezal de lectura/escritura que flota sobre una delgada lámina de aire generada por la rotación de los discos.



Figura 8-17. Disco duro.

El disco duro presenta otros elementos además del soporte magnético:

1. **Código de canal.** Los discos duros incluyen técnicas FEC para obtener tasas de error de bit muy bajas, del orden de 10^{-14} . Los discos duros trabajan al límite de Nyquist, es decir, con $R_b = 2B$. Como no existe desgaste de las cabezas, el valor de B (ancho de banda) no disminuye con el uso, y el régimen binario se mantiene.
2. **Buffer o Cache.** Cuando se reproduce régimen binario variable, bien de canal o bien de usuario, se convierte en régimen binario fijo de usuario gracias a la memoria buffer. La memoria buffer también sirve para grabar a una velocidad adecuada al disco todos los datos.
3. **Interfaz al bus.** Es la circuitería electrónica que conecta el buffer del disco con el bus exterior de conexión de discos, luego es específico de cada bus. Generalmente se usan buses en paralelo por lo que el régimen binario en el interfaz se especifica en Bytes en vez de en bits. La interfaz debe ser capaz de manejar el régimen binario del bus y el del disco, que son por lo general:
 - Rb sostenido: el que se puede asegurar de forma continuada (para vídeo).
 - Rb de pico: en casos de ficheros cortos que estén disponibles íntegros en el buffer, estos pueden transferirse al máximo, que está limitado sólo por la interfaz.

Durante la grabación, el régimen binario grabado en cada pista queda determinado por el controlador del disco que controla la transferencia entre el buffer y las cabezas. Durante la reproducción, la propia señal previamente grabada en el disco determina el régimen binario de lectura. Según se controle la grabación se puede tener:

1. **Modo fijo:** se graban todas las pistas al mismo R_b , que deberá ser igual al valor de la pista con menor régimen binario. Por ello, no se aprovecha la mayor capacidad de las pistas exteriores, aunque se facilitan las aplicaciones de grabación y reproducción sostenidas de fragmentos muy largos.
2. **Modo variable:** cada pista se graba a su régimen binario máximo. El R_b de canal es variable y el disco se aprovecha al máximo. Sin embargo las transferencias sostenidas de larga duración son problemáticas.

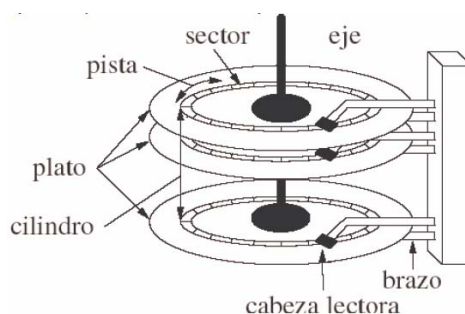


Figura 8-18. Estructura interna de un disco duro.

Dentro de un disco duro hay uno o varios discos (de aluminio o cristal) concéntricos llamados platos (normalmente entre 2 y 4, aunque pueden ser hasta 6 o 7 según el modelo), y que giran todos a la vez sobre el mismo eje, al que están unidos. El cabezal (dispositivo de lectura y escritura) está formado por un conjunto de brazos paralelos a los platos, alineados verticalmente y que también se desplazan de forma simultánea, en cuya punta están las cabezas de lectura/escritura. Por norma general hay una cabeza de lectura/escritura para cada superficie de cada plato. Los cabezales pueden moverse hacia el interior o el exterior de los platos, lo cual combinado con la rotación de los mismos permite que los cabezales puedan alcanzar cualquier posición de la superficie de los platos.

Un disco duro consta de los siguientes elementos en su estructura:

- **Plato:** cada uno de los discos que hay dentro del disco duro.
- **Cara:** cada uno de los dos lados de un plato.
- **Cabeza:** número de cabezales.
- **Pistas:** a nivel lógico, es una circunferencia dentro de una cara; la pista 0 está en el borde exterior.

- **Cilindro:** conjunto de varias pistas; son todas las circunferencias que están alineadas verticalmente (una de cada cara).
- **Sector:** cada una de las divisiones de una pista. El tamaño del sector no es fijo, siendo el estándar actual 512 bytes, aunque se puede llegar hasta los 4 KiB.

El disco no para de girar a velocidad angular constante (por ejemplo 5600 rpm, 7200 rpm, 10000 rpm o 15000 rpm). Durante cada vuelta, la cabeza permanece parada, por lo que la huella creada es una pista circular. Un disco con varias cabezas no las usa simultáneamente para grabar/reproducir varios canales internos en paralelo, sino que las conmuta secuencialmente manejando un solo canal interno.

La velocidad cabeza-pista es igual a la lineal de cada pista pues la cabeza permanece quieta, cosa que no ocurría en los magnetoscopios. La velocidad por tanto vendrá dada por la expresión (8-3):

$$V_{cp} = \omega r_p = 2\pi V_g r_p \text{ rad/s. (8-3)}$$

Donde ω y V_g son la velocidad de giro de los discos en rad/s y rpm respectivamente; y r_p es el radio de cada pista. El radio de cada pista es la separación de ésta respecto al centro del disco. Cada pista tendrá una velocidad cabeza-pista distinta y por lo tanto un ancho de banda calculado con la expresión (8-4):

$$f_{max} = \frac{V_{cp}}{2e} = \frac{\pi V_g r_p}{e} \text{ Hz. (8-4)}$$

Por lo tanto, el régimen binario máximo manejable es distinto en cada pista, siendo más alto en las pistas externas y más bajo en las internas.

Haciendo trabajar de forma conjunta a varios discos se pueden obtener mejoras en fiabilidad (mediante Técnicas de Redundancia) y en rendimiento (mediante Paralelismo).

Para incrementar la fiabilidad de un sistema de almacenamiento masivo se usa replicación de contenido. La técnica más simple es la del espejo (*mirroring*) y consiste en ejecutar las operaciones de E/S en dos dispositivos idénticos y que por tanto tienen el mismo contenido, con esto cuando un disco sufre una avería, se usa el segundo disco mientras se repara el primero.

Para incrementar el rendimiento de un sistema de almacenamiento masivo se usa paralelismo. Con esta técnica si el tiempo de acceso a un dato de un disco es x ms y se dispone de n discos, es posible acceder a n datos en x ms, luego se mejora la velocidad de transferencia para múltiples discos.

8.4.2.1 Sistema de conexión en paralelo

La conexión de varios discos en paralelo conectados a un bus común permite acercarse a tres objetivos:

1. **Aumento del régimen binario asignable al manejo de una señal.** La señal con un régimen binario alto se divide en ‘canales’ de menor régimen binario y cada uno se graba en un disco distinto. Todos los discos trabajan de forma sincronizada gracias al controlador, que a su vez se encarga del demultiplexado/multiplexado de canales.
2. **Aumento del número de señales manejadas simultáneamente.** Cada una de las señales deberá tener un régimen binario suficientemente bajo como para ser manejada por un solo disco duro. El mismo proceso de multiplexación permite enviar cada señal a su disco. Cada una de las señales debe llegar al controlador debidamente multiplexada en tiempo.
3. **Aumento de la capacidad de almacenamiento o duración de la grabación.** Los discos conectados en paralelo a un bus común también permiten un uso secuencial, no simultáneo, de forma que en cada momento de tiempo sólo se usa uno de ellos. El bus funciona al régimen binario de un disco. Lo que se consigue es tener mayor capacidad de datos.

Los dos primeros objetivos exigen que el bus de conexión trabaje con un régimen binario superior al de un disco al multiplexar el envío de datos a varios discos simultáneamente. Para el tercero se usan los discos secuencialmente y no simultáneamente por lo no incrementa la carga del bus.

Las controladoras de discos, actúan como puentes entre el bus de discos y el bus principal del ordenador. Sin embargo, el control de agrupaciones paralelo puede suponer la ejecución de funciones específicas, lo que conlleva un software específico e incluso un hardware añadido. Así, son funciones específicas la separación de datos en canales y la generación de datos sincronización y de paridad para proteger el funcionamiento de la agrupación. La primera y segunda puede resolverse sólo con el software, pero la tercera suele requerir circuitería adicional.

En cuanto a los buses de discos, los más comunes son el SCSI y ATA, también llamados IDE y pueden estar implementados como buses paralelo o serie.

8.4.2.2 Sistemas RAID

Las agrupaciones o arrays de discos según RAID (*Redundant Array of Independent Disks*) manejan los siguientes conceptos:

1. *Stripping*. Es la agrupación y formateo de 2 o más discos para formar una única unidad lógica.
2. *Chunk*. Recibe ese nombre cada uno de los fragmentos de cada uno de los trenes de datos que se graban en un disco del array. Pueden corresponderse a cada uno de los canales obtenidos por demultiplexación de la señal original según la configuración y uso del RAID.
3. *Parity*. Es la información redundante que se usa para corregir errores debidos a defectos del funcionamiento del array, y así recuperar la señal original con más fiabilidad.
4. *Transfer Rate Intensive*. Es el modo de funcionamiento del array en el que el conjunto de discos funciona en paralelo para manejar una única señal correspondiente a una petición. Si son muchos discos, cada *chunk* es pequeño al haber mucha demultiplexación.
5. *Request Rate Intensive*. Es el modo de funcionamiento del array en el que cada disco del conjunto maneja señal de petición concreta. El conjunto maneja varias señales o peticiones simultáneamente. El régimen binario requerido de cada petición es pequeño. El *chunk* suele ser mayor.

Los niveles o estructuras RAID, de uso habitual en vídeo son:

1. RAID-0. Se dividen los datos entre los discos sin incluir redundancias. Es un simple *striping* que ejecuta y controla con el software del sistema operativo. No es aconsejable en equipos profesionales de uso crítico.

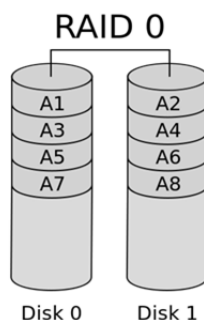


Figura 8-19. Esquema de funcionamiento del RAID-0.

2. RAID-1. Llamado *Mirroring*. Consiste en grabar toda la información por duplicado en discos distintos del array. Se usa en servidores de emisión donde la máxima seguridad es un requisito.

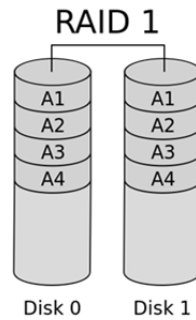


Figura 8-20. Esquema de funcionamiento del RAID-1.

3. RAID-3. Requiere la generación de datos de paridad que se graban en un disco específico del array. Optimizado para modos *Transfer Rate Intensive*.
4. RAID-5. Requiere la generación de datos de paridad que se graban repartidos en los discos del array. Optimizado para modos *Request Rate Intensive*.

RAID-3

Se usa $(n+1)$ discos para cada unidad lógica, de modo que los datos se dividen entre n discos y la redundancia calculada se graba en el otro disco. El controlador está optimizado para aplicaciones *Transfer Rate Intensive*, manejando *chunks* pequeños y asegurando que los giros de los discos van sincronizados. Sólo responde a una petición simultánea y por tanto su rendimiento es pobre ante el manejo de operaciones I/O (*input/output*) aleatorias.

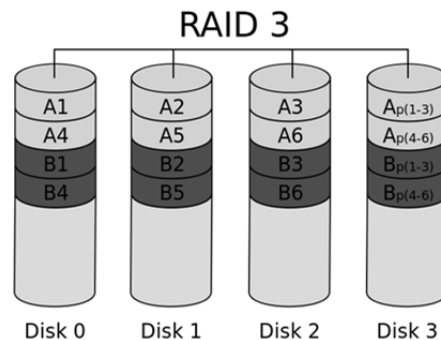


Figura 8-21. Esquema de funcionamiento del RAID-3.

Este comportamiento hace de RAID-3 una solución muy buena para sistemas de edición no lineal en lo que se graban señales de una en una y se reproducen pocas señales simultáneas y casi siempre sólo una, salvo durante las transiciones. También es buena solución para servidores que manejan varias señales pero cada una de ellas usando una unidad lógica RAID-3 distinta. La tarjeta controladora debe llevar la circuitería que genera los datos redundantes, que funciona igual que un FEC. Estos datos se graban en un único disco dedicado. Durante la reproducción todos los datos y las redundancias se analizan en la circuitería de la tarjeta controladora y se corrigen los errores que se detecten.

Una situación límite para la corrección de errores es que el RAID corrija todos los errores, y por lo tanto reproduzca correctamente, ante la pérdida de uno de los discos duros del array. Esto es posible con el RAID-3 pero requiere de una alta cantidad de discos, lo que supondría a su vez la infrautilización de estos.

RAID-5

Se usan n discos para cada unidad lógica. La redundancia y la señal se reparten entre los n discos. En lugar de almacenar el bloque de paridad en un único disco, se almacena de forma distribuida con los discos de datos. Para un determinado bloque, un disco almacena la paridad y el resto los datos. Para otro bloque, la paridad está almacenada en un disco diferente. Un bloque de paridad no puede almacenar paridad para la información

almacenada en el mismo disco. Se consigue distribuir la carga de operaciones de forma homogénea entre todos los discos.

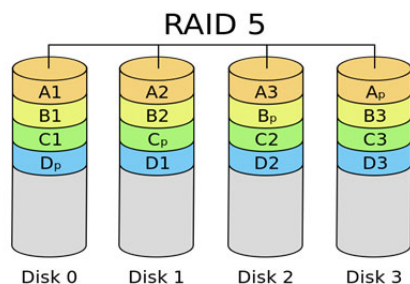


Figura 8-22. Esquema de funcionamiento del RAID-5.

RAID-0+1

Combina las prestaciones de RAID 0 y RAID 1. Es un sistema implementado como una tabla distribuida por bloques. Cada dispositivo que contiene un bloque, a su vez contiene un dispositivo redundante (RAID 0). Si un disco falla, anula la distribución, y se pasa a utilizar la distribución en el espejo (RAID 1).

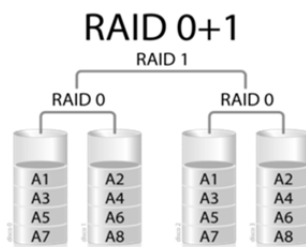


Figura 8-23. Esquema de funcionamiento del RAID-0+1.

8.5 Formatos de archivado

Los formatos de archivo especifican la estructura específica de los datos multimedia en un archivo o dispositivo de almacenamiento. El formato se puede “deducir” a partir de las letras que componen la extensión del nombre del archivo, por ejemplo en mivideo.wmv se presupone un formato WMV.

8.5.1 Códec

Abreviatura de CODificador-DECodificador. Describe una especificación de software, hardware o combinación de ambos, capaz de transformar un archivo con un flujo de datos (*stream*) o una señal. En general un códec es un algoritmo de software que transforma datos de un formato a otro. Un códec puede codificar un *stream* o una señal para transmitir, almacenar o cifrar y posteriormente permite su recuperación y/o descifrado para reproducirlo o manejarlo en el formato apropiado.

La mayor parte de los códec persiguen reducir el tamaño del archivo y por eso suelen provocar pérdida de información, aunque también existen códec sin pérdidas (*lossless*). En general un archivo multimedia contiene datos de audio y vídeo, además de otros de sincronización, metadatos, etc. y dado que cada uno de esos flujos va a ser tratado mediante diferentes aplicaciones y equipos, se encapsulan juntos dentro de un archivo de “contenedor” tipo .avi, .mov, .mp4, etc. Existen cientos de Códec diferentes, por lo que en general un usuario necesitará combinaciones de varios de ellos para poder utilizar su reproductor multimedia con diferentes archivos. Hay Códec para la compresión de audio y video, para poder reproducir archivos en modo “*streaming*” por Internet, videoconferencia, etc.

A nivel de usuario una “solución fácil” al problema de la multitud de códec es usar un paquete de Códec, que simplemente es una colección de Códec agrupados en un único fichero de gran tamaño. El más conocido es K-

Lite Códec Pack. Una confusión muy común es creer que AVI es un códec, cuando realmente es un formato de fichero contenedor Multimedia. Dado que existen cientos de Códec compatibles con AVI, en ocasiones puede ser complicado saber cuál es el que necesitamos para un archivo específico. El diseño del códec ha de tener en cuenta el tipo de información a codificar, por eso existen códec diferentes para imagen fija/video o voz/música.

Existen dos tipos de códec:

- **lossless** (Sin pérdidas): Se puede recuperar exactamente el archivo original a partir del codificado.
- **lossy** (Con pérdidas): El archivo decodificado es una aproximación del original, que depende del nivel de compresión usado.

Algunos ejemplos de formatos de audio:

AAC ("Audio Advanced Coding"): Codificado Avanzado de Audio.

AIFF ("Audio Interchange File Format"): Codesarrollado por Apple.

CDA ("Compact Disc Audio"): Presente en los CD-DA.

MP3 ("Moving Picture expert group - layer 3"): Audio comprimido con pérdida.

REC ("RECORD"): Audio de grabadoras de voz digitales.

WAV ("WAVEform audio format"): Formato de audio de Microsoft® e IBM®.

WMA ("Windows Media Audio"): Alternativa de Microsoft a MP3.

Algunos ejemplos de formatos de video:

MPEG ("Moving Picture Expert Group")

WMV ("Windows Media Video")

3GP ("3a Generation Partnership")

AVI ("Audio Video Interleave")

VCD ("Video Compact Disc")

SVCD ("Super Video Compact Disc")

MP4 ("Moving Picture expert group - 4 part 14")

Algunos ejemplos de formatos de imágenes:

JPEG ("Joint Photographic Expert Group")

PNG ("Portable Network Graphics")

GIF ("Graphics Interchange Format")

BMP ("BitMaP")

TIFF ("Tagged Image")

8.5.2 Formatos de audio

Un formato de archivo de audio es un contenedor multimedia que guarda una grabación de audio. El procedimiento de reproducción y su uso en un sistema de ficheros variará dependiendo del formato, como también lo hará el modo en que se almacenan los datos dentro del archivo. Los datos pueden almacenarse sin comprimir o comprimidos para reducir el tamaño del formato.

Hay que tener especial atención a la hora de diferenciar códec con el formato de un archivo. El códec es el método de codificación de un archivo que en sí ya tiene un formato propio. Existen diferentes tipos de formato según la compresión del audio.

Se pueden encontrar formatos sin pérdida, por los cuales se obtiene la señal sin ningún tipo de procesado para reducir su tamaño (compresión); y formatos con compresión, que a su vez se dividen en compresión sin pérdidas (*loseless*), en los que se emplean algoritmos que no eliminan información original para reducir el tamaño de un archivo; y en compresión con pérdidas (*lossy*), en los que se elimina información original de la señal sin reducir notablemente su calidad para conseguir reducir el tamaño de los archivos.

8.5.2.1 Formatos sin pérdida

.WAV

De *wave* en inglés. Es un formato de audio digital normalmente sin compresión de datos, desarrollado y propiedad de Microsoft y de IBM que se utiliza para almacenar sonidos en el PC, admite archivos mono y estéreo a diversas resoluciones y velocidades de muestreo, su extensión es .wav.

Es una variante del formato RIFF (*Resource Interchange File Format*, formato de fichero para intercambio de recursos), método para almacenamiento en "paquetes", y relativamente parecido al IFF y al formato AIFF usado por Macintosh. El formato toma en cuenta algunas peculiaridades de la CPU Intel, y es el formato principal usado por Windows.

A pesar de que el formato WAV es compatible con casi cualquier códec de audio, se utiliza principalmente con el formato PCM (no comprimido) y, al no tener pérdida de calidad, es adecuado para uso profesional. Para tener calidad CD de audio se necesita que el sonido se grabe a 44100 Hz y a 16 bits. Por cada minuto de grabación de sonido se consumen unos 10 megabytes de espacio en disco. Una de sus grandes limitaciones es que solo se pueden grabar archivos de 4 gigabytes como máximo, lo cual equivale aproximadamente a 6,6 horas en calidad de CD de audio. Es una limitación propia del formato, independientemente de que el sistema operativo donde se utilice sea MS Windows u otro distinto, y se debe a que en la cabecera del fichero se indica la longitud del mismo con un número entero de 32 bits, lo que limita el tamaño del fichero a un máximo de 4294967295 bytes (≈ 4 GB)

En Internet no es popular, fundamentalmente porque los archivos sin compresión son muy grandes. Son más frecuentes los formatos comprimidos con pérdida, como el MP3 o el Ogg Vorbis. Como éstos son más pequeños, la transferencia a través de Internet es mucho más rápida. Además, existen códecs de compresión sin pérdida más eficaces, como Apple Lossless o FLAC.

.AIFF

Audio Interchange File Format. Es un estándar de formato de audio usado para almacenar datos de sonido en computadoras personales. El formato fue codesarrollado por Apple Inc. en 1988 basado en el IFF (*Interchange File Format*) de Electronic Arts, usado internacionalmente en las computadoras *Amiga* y actualmente es muy utilizado en las computadoras Apple Macintosh. AIFF también es el utilizado por Silicon Graphics Incorporated.

Los datos de audio en el estándar AIFF no están comprimidos, almacenándose los datos en *big-endian* y emplea una modulación por impulsos codificados (PCM). También hay una variante del estándar donde sí que existe compresión, conocida como **AIFF-C** o **AIFC**, con varios códecs definidos.

El estándar AIFF es uno de los formatos líderes, junto a SDII y WAV, usados a nivel profesional para aplicaciones de audio ya que, a diferencia del conocido formato con pérdidas MP3, éste formato está comprimido sin ninguna pérdida, lo que ayuda a un rápido procesado de la señal pero con la desventaja del gran

espacio en disco que supone: alrededor de 10MB para un minuto de audio estéreo con una frecuencia de muestreo de 44.1kHz y 16 bits. Además el estándar da soporte a bucles para notas musicales para uso de aplicaciones musicales o *samplers*. Las extensiones de archivo para el formato estándar AIFF son **.aiff** o **.aif**. Para las variantes comprimidas se supone que es **.aifc**, pero las anteriores también son aceptadas por las aplicaciones que soportan este tipo de formato.

8.5.2.2 Formatos con compresión sin pérdida

Necesitan más tiempo de procesado que los formatos no comprimidos, aunque por otro lado tienen mayor eficiencia en consumo de soporte de almacenamiento, frente a los formatos de audio sin comprimir que codificaban audio y silencio con el mismo nº de bits por unidad de tiempo, codificar un minuto de silencio en un formato sin comprimir producía un archivo del mismo tamaño que codificar un minuto de música. Es este tipo de formatos la música ocupa un tamaño ligeramente inferior, mientras que el silencio no ocupa prácticamente nada. Presentan ratio de compresión de aproximadamente 2:1 respecto a sin pérdida y los desarrollos actuales se centran en la reducción del tiempo de procesado manteniendo el ratio de compresión.

.WMA

Windows Media Audio es una tecnología de compresión de audio desarrollada por Microsoft. En nombre puede usarse para referirse al formato de archivo de audio o al códec de audio. Es software propietario que forma parte de la suite Windows Media.

WMA consiste de cuatro códec distintos. El códec WMA original, conocido simplemente como WMA, fue concebido como competidor al MP3 y al RealAudio. *WMA Pro*, un códec más moderno y avanzado, soporta audio *surround* y de alta resolución. El formato en sí tiene compresión con pérdida pero también está disponible la versión *WMA Lossless*, el cual comprime audio sin perder definición. Existe otra variación llamada *WMA Voice*, enfocada en contenido hablado, aplica compresión y está diseñado para tasas de bits muy bajas.

Microsoft afirma que el audio codificado con WMA provee una mejor calidad de sonido que MP3 a la misma tasa de bits; Microsoft también afirma que el audio codificado con WMA con tasas de bits menores tiene mejor calidad que MP3 codificado con tasas de bits superiores. Pruebas de escucha doble ciego (sin influencia de ningún tipo) con otros códec de audio con pérdida muestran resultados diversos, algunos desmintiendo las afirmaciones de Microsoft, otros comprobándolas. Una prueba independiente hecha en mayo de 2004 con una tasa de bits de 128 kbit/s mostró que WMA es equivalente al LAME MP3; inferior al AAC y al Vorbis; y superior al ATRAC3.

Permiten hasta tasas de 96 Khz con resolución de 24 bits y hasta 6 canales (válido para audio 5.1). El ratio de compresión varía (dependiendo del audio que se comprima) entre 1.7:1 (mucha información) y 3:1 (poca información ó mucho silencio).

8.5.2.3 Formatos con compresión con pérdida

La codificación comprime los datos de audio descartando partes de ellos. El proceso intenta minimizar cantidad de datos que mantiene el archivo reduciendo su volumen a costa de la reducción de calidad. Se basa en el funcionamiento de oído humano: sólo pierde las frecuencias “poco o nada” audibles por el oído humano, conservando gran parte de su calidad. Ejemplos de este tipo de formatos son el MP3, AAC y WMA (con pérdida).

.MP3: MPEG-1 Audio Layer III o MPEG-2 Audio Layer III, más comúnmente conocido como MP3, es un formato de compresión de audio digital patentado que usa un algoritmo con pérdida para conseguir un menor

tamaño de archivo. Es un formato de audio común usado para música tanto en ordenadores como en reproductores de audio portátil.

Los archivos MPEG-1 corresponden a las velocidades de muestreo de 32, 44.1 y 48 kHz.

Los archivos MPEG-2 corresponden a las velocidades de muestreo de 16, 22.05 y 24 kHz.

MP3 fue desarrollado por el Moving Picture Experts Group (MPEG) para formar parte del estándar MPEG-1 y del posterior y más extendido MPEG-2. Un MP3 creado usando una compresión de 128kbit/s tendrá un tamaño de aproximadamente unas 11 veces menor que su homónimo en CD. Un MP3 también puede comprimirse usando una mayor o menor tasa de bits por segundo, resultando directamente en su mayor o menor calidad de audio final, así como en el tamaño del archivo resultante.

Estudios demuestran que la distorsión generada es imperceptible para el oído experimentado en un ambiente óptimo desde los 192 kbps y en condiciones normales. Para el oído no experimentado, o común, con 128 kbps o hasta 96 kbps basta para que se oiga "bien" (a menos que se posea un equipo de audio de alta calidad donde se nota excesivamente la falta de graves y se destaca el sonido de "fritura" en los agudos). En personas que escuchan mucha música o que tienen experiencia en la parte auditiva, desde 192 o 256 kbps basta para oír bien. La música que circula por Internet, en su mayoría, está codificada entre 128 y 192 kbps.

.AAC

Advanced Audio Coding, es un formato informático de señal digital de audio basado en un algoritmo de compresión con pérdida, un proceso por el que se eliminan algunos de los datos de audio para poder obtener el mayor grado de compresión posible, resultando en un archivo de salida que suena lo más parecido posible al original.

El formato AAC corresponde al estándar internacional "ISO/IEC 13818-7" como una extensión de MPEG-2: un estándar creado por MPEG (Moving Pictures Expert Group).

Debido a su excepcional rendimiento y la calidad, la codificación de audio avanzada (AAC) se encuentra en el núcleo del MPEG-4, 3GPP y 3GPP2, y es el códec de audio de elección para Internet, conexiones inalámbricas y de radio difusión digital.

Este formato AAC ha sido elegido por Apple como formato principal para los iPods y para su software iTunes. También es utilizado en otras aplicaciones por Ahead Nero, Winamp y Nintendo DSi. El formato AAC a diferencia del formato OGG, permite incluir legalmente la protección de los derechos de autor, aquellos archivos de audio sin autorización, que tengan protección anticopia no funcionarán en AAC.

El AAC utiliza una frecuencia de bits variable (VBR), un método de codificación que adapta el número de bits utilizados por segundo para codificar datos de audio, en función de la complejidad de la transmisión de audio en un momento determinado.

AAC es un algoritmo de codificación de banda ancha de audio que tiene un rendimiento superior al del MP3, que produce una mejor calidad en archivos pequeños y requiere menos recursos del sistema para codificar y decodificar.

Este códec está orientado a usos de banda ancha y se basa en la eliminación de redundancias de la señal acústica, así como en compresión mediante la transformada de coseno discreta modificada (MDCT), muy parecido al del MP3. No compatible con MPEG-1.

Frecuencia de muestreo: 96 KHz, 88.2 KHz, 48 KHz, 44.1 KHz, 24 KHz, 22.05 KHz, 16 KHz.

Máxima calidad entre 320 y 448 kbps (5 canales) y entre 128 y 192 (2 canales)

Máximo *Bit Rate* Soportado: De 12 kbps (En Perfil HE-AACv2) hasta 448 kbps (En Perfil AAC-LC)

Tres opciones:

- Máxima calidad (resolución a 23.43 Hz y 2.6 ms)
- No predicción
- Frecuencias de muestreo escalables

Tres tipos de perfiles:

- AAC-LC: De 16 KBps a 448 KBps (Perfil más eficiente en calidad de audio).
- HE-AAC: De 16 KBps a 128 KBps.
- HE-AACv2: De 12 KBps a 56 KBps.

8.5.2.4 Protocolo MIDI

MIDI son las siglas de la (Interfaz Digital de Instrumentos Musicales). Se trata de un protocolo de comunicación serial estándar que permite a los computadores, sintetizadores, secuenciadores, controladores y otros dispositivos musicales electrónicos comunicarse y compartir información para la generación de sonidos.

El estándar MIDI fue inicialmente propuesto en un documento dirigido a la Audio Engineering Society por Dave Smith, presidente de la compañía Sequential Circuits en 1981. La primera especificación MIDI se publicó en agosto de 1983.

Cabe aclarar que MIDI no transmite señales de audio, sino datos de eventos y mensajes controladores que se pueden interpretar de manera arbitraria, de acuerdo con la programación del dispositivo que los recibe. Es decir, MIDI es una especie de "partitura" que contiene las instrucciones en valores numéricos (0-127) sobre cuándo generar cada nota de sonido y las características que debe tener; el aparato al que se envíe dicha partitura la transformará en música completamente audible.

En la actualidad la gran mayoría de los creadores musicales utilizan el protocolo MIDI a fin de llevar a cabo la edición de partituras y la instrumentación previa a la grabación con instrumentos reales. Sin embargo, la perfección adquirida por los sintetizadores en la actualidad lleva a la utilización de forma directa en las grabaciones de los sonidos resultantes del envío de la partitura electrónica a dichos sintetizadores de última generación.

Buena parte de los dispositivos MIDI son capaces de enviar y recibir información, pero desempeñan un papel diferente dependiendo de si están recibiendo o enviando información; también depende de la configuración del programa o programas que pueden usar dicho dispositivo. El que envía los mensajes se denomina Maestro (del inglés *master*, o 'amo') y el que responde a esa información, Esclavo (*slave*).

Dado que solo contiene instrucciones, los ficheros MIDI files pueden ser muy pequeños. Este formato es compatible con diferentes plataformas de software y navegadores de Internet. Los ficheros MIDI tienen la extensión .mid o .midi.

8.5.3 Formatos de vídeo

Los videos digitales se pueden guardar en archivos de distintos formatos. Cada uno se corresponde con una extensión específica del archivo que lo contiene. Existen muchos tipos de formatos de video. Aquí se citan algunos de los más utilizados. Asimismo cada tipo de archivo admite en cada momento un códec de compresión distinto. Se obvia el que el vídeo lleva consigo audio, luego los formatos de vídeo soportan ambos tipos de información.

A continuación se va a hacer una introducción a los formatos más extendidos.

.AVI

Audio Video Interleave, es un formato contenedor de audio y video lanzado por Microsoft en 1992. El formato avi permite almacenar simultáneamente un flujo de datos de video y varios flujos de audio. El formato concreto de estos flujos no es objeto del formato AVI y es interpretado por un programa externo denominado códec. Es decir, el audio y el video contenidos en el AVI pueden estar en cualquier formato (AC3/DivX, u MP3/Xvid, entre otros). Por eso se le considera un formato contenedor.

Para que todos los flujos puedan ser reproducidos simultáneamente es necesario que se almacenen de manera intercalada (*interleave*). De esta manera, cada fragmento de archivo tiene suficiente información como para reproducir unos pocos fotogramas junto con el sonido correspondiente.

Obsérvese que el formato AVI admite varios flujos de datos de audio, lo que en la práctica significa que puede contener varias bandas sonoras en varios idiomas. Es el reproductor multimedia quien decide cuál de estos flujos debe ser reproducido, según las preferencias del usuario.

Los archivos AVI se dividen en fragmentos bien diferenciados denominados *chunks*. Cada *chunk* tiene asociado un identificador denominado *etiqueta FourCC*. El primer fragmento se denomina *cabecera* y su papel es describir *meta-información* respecto al archivo, por ejemplo, las dimensiones de la imagen y la velocidad en fotogramas por segundo. El segundo *chunk* contiene los riuos flujos entrelazados de audio y video. Opcionalmente, puede existir un tercer *chunk* que actúa a modo de índice para el resto de *chunks*.

Para reproducir un archivo AVI es necesario lo siguiente:

- Un reproductor de video capaz de interpretar el formato AVI.
- El *códec* de video para interpretar el flujo de video.
- El *códec* de audio para interpretar el flujo de audio.

La etiqueta FourCC permite identificar el *códec* necesario para interpretar un flujo de audio o video. Cada códec tiene asociados el conjunto de etiquetas que es capaz de reproducir. De esta manera, el reproductor de video es capaz de elegir el códec pertinente sin intervención del usuario.

El reproductor consecutivamente lee fragmentos del archivo de audio y música y sonidos AVI. Después separa cada uno de los flujos de audio y video que se encuentran intercalados en el archivo. Cada uno de estos flujos, una vez separados, se almacena en un *buffer* de memoria y se pasan al *códec* correspondiente. El códec de video devuelve otro *buffer* que contiene cada uno de los fotogramas a reproducir. El códec de audio retorna otro *buffer* con la muestra digital de sonido a reproducir. Con esta información, el reproductor solamente tiene que sincronizar los fotogramas y el sonido y reproducirlos a la velocidad adecuada..

En el sistema operativo Mac OS es perfectamente posible visualizar archivos AVI, siempre que los códecs utilizados estén soportados por Quicktime, bien directamente o a través de plugins. Existen otros reproductores de archivos multimedia para esta plataforma que también permiten visualizar correctamente archivos AVI.

-.mov: Es el formato de video y audio desarrollado por Apple. Utiliza un códec propio que evoluciona en versiones con bastante rapidez. Este tipo de archivos también pueden tener extensión *.QT. Se recomienda utilizar el reproductor de QuickTime. Existe una versión gratuita del mismo que se puede descargar de Internet. Es ideal para publicar videos en Internet por su razonable calidad/peso. Admite *streaming*.

-.mpg: MPEG-1 es el nombre de un grupo de estándares de codificación de audio y vídeo normalizados por el grupo MPEG (*Moving Pictures Experts Group*). MPEG-1 vídeo se utiliza en el formato Video CD. La calidad de salida con la tasa de compresión usual usada en VCD es similar a la de un cassette vídeo VHS doméstico. MPEG-1 está conformado por diferentes partes:

1. Sincronización y transmisión simultánea de vídeo y audio.
2. Códec de compresión para señales de vídeo no entrelazadas (progresivas).
3. Códec de compresión para señales de audio con control sobre la tasa de compresión. El estándar define tres *capas* (layers en inglés), o niveles de complejidad de la codificación de audio MPEG.
 1. MP1 o MPEG-1 Parte 3 Capa 1 (MPEG-1 Audio Layer 1)
 2. MP2 o MPEG-1 Parte 3 Capa 2 (MPEG-1 Audio Layer 2)
 3. MP3 o MPEG-1 Parte 3 Capa 3 (MPEG-1 Audio Layer 3)
4. Procedimientos para verificar la conformidad.
5. Software de referencia.

Desde su creación se han definido el MPEG-1 (CD-ROM y Vídeo CD), el MPEG-2, (DVD de Vídeo y TV Digital) y **MPEG-4**, que se emplea para transmitir vídeo e imágenes en ancho de banda reducido, además de otros usos. Formato multiplataforma y multinavegador. Es un formato adecuado para distribuir en la Web. MPEG4 es la base de formatos como **divx**, **xvid** o **H264**. Los archivos MPEG usan extensión .mpg o .mpeg.

-.wmv: es un nombre genérico que se da al conjunto de algoritmos de compresión ubicados en el set propietario de tecnologías de vídeo desarrolladas por Microsoft, que forma parte del *framework* Windows Media.

WMV no se construye sólo con tecnología interna de Microsoft. Desde la versión 7 (WMV1), Microsoft ha utilizado su propia versión no estandarizada de MPEG-4. El vídeo a menudo se combina con sonido en formato Windows Media Audio.

El formato WMV es reproducido por una amplia gama de reproductores, como BS.Player, MPlayer o Windows Media Player, el último sólo disponible en plataformas Windows y Macintosh (sin compatibilidad completa). En el caso de reproductores ajenos a Microsoft, como por ejemplo el citado MPlayer, es frecuente utilizar una implementación alternativa de los formatos, como por ejemplo la de FFmpeg.

El vídeo WMV se empaqueta normalmente en algún contenedor multimedia, como pueden ser AVI o ASF. Los ficheros resultantes reciben la extensión .avi si el contenedor es de este tipo, .wmv si es un fichero de sólo video (.wma sería el equivalente para sonido) o .asf si se trata de un contenedor ASF, con contenido de audio y vídeo.

El formato WMV incluye ciertas características relativas a la utilización de DRM (gestión digital de derechos). Sin embargo, estas características pueden eliminarse con utilidades como Sidda, DRMCreep, drmdbg o drm2wmv.

La versión 9 del formato utilizado por Microsoft fue remitida a la *Society Motion Picture and Televisión Engineers* (SMPTE) para su aprobación como estándar, bajo el nombre en clave VC-1. Este códec es también utilizado en la distribución de vídeo de alta definición sobre DVD estándar en un formato que Microsoft ha denominado WMV HD, el cual puede ser reproducido tanto en ordenadores como en reproductores de sobremesa.

Una primera versión del estándar fue publicada por la SMPTE en septiembre de 2005, incluyéndose información para la implementación de un decodificador y secuencias de prueba.

-.flv: Flash Video (FLV) es un formato contenedor propietario usado para transmitir video por Internet usando Adobe Flash Player (anteriormente conocido como Macromedia Flash Player), desde la versión 6 a la 10. Los contenidos FLV pueden ser incrustados dentro de archivos SWF. Entre los sitios más notables que utilizan el formato FLV se encuentran YouTube, Google Video, Reuters.com, Yahoo! Video y MySpace.

Flash Video puede ser visto en la mayoría de los sistemas operativos, mediante Adobe Flash Player, el plugin extensamente disponible para navegadores web, o de otros programas de terceros como MPlayer, VLC media player, o cualquier reproductor que use filtros DirectShow (tales como Media Player Classic, Windows Media Player, y Windows Media Center) cuando el filtro ffdshow está instalado.

8.5.4 Formatos de archivado de imagen

Existen multitud de formatos gráficos para guardar imágenes (JPG, PNG, GIF, BMP...). Cada uno de ellos tiene sus ventajas e inconvenientes, y dependiendo de una serie de criterios y el uso que se le quiera dar a la imagen, es mejor utilizar uno u otro. En este manual, vamos a ver, desde un punto de vista algo técnico (pero comprensible) cuales son las características de varios formatos de imagen relevantes, así como varios programas para optimizarlos.

Existen ciertas diferencias entre los formatos en los siguientes aspectos:

-Niveles de profundidad admitidos:

1 bit → Imágenes en blanco y negro

1 byte → Escala de grises o paleta de 256 colores

3 bytes → Modelo RGB

-Tipo de compresión:

Sin pérdida: RLE, LZW

Con pérdida: mediante FFT, DCT, wavelets

-Otras características:

Posibilidad de definir de transparencias.

Diferentes imágenes en un mismo archivo (animaciones).

Según la aplicación será más adecuado uno u otro formato.

.BMP

Windows bitmap es el formato propio del programa, que viene con el sistema operativo Windows. Puede guardar imágenes de 24 bits (16,7 millones de colores), 8 bits (256 colores) y menos. Puede darse a estos archivos una compresión sin pérdida de calidad: la compresión RLE (*Run-length encoding*).

Los archivos con extensión **.BMP**, en los sistemas operativos Windows, representan la sigla BitMaP (o también Bit Mapped Picture), o sea mapa de bits. Los archivos de mapas de bits se componen de direcciones asociadas a códigos de color, uno para cada cuadro en una matriz de píxeles tal como se esquematizaría un dibujo de "colorea los cuadros" para niños pequeños. Normalmente, se caracterizan por ser muy poco eficientes en su uso de espacio en disco, pero pueden mostrar un buen nivel de calidad. A diferencia de los gráficos vectoriales, al ser reescalados a un tamaño mayor, pierden calidad. Otra desventaja de los archivos BMP es que no son utilizables en páginas web debido a su gran tamaño en relación a su resolución.

Dependiendo de la profundidad de color que tenga la imagen cada píxel puede ocupar 1 o varios bytes. Generalmente se suelen transformar en otros formatos, como JPEG (fotografías), GIF o PNG (dibujos y esquemas), los cuales utilizan otros algoritmos para conseguir una mayor compresión (menor tamaño del archivo).

Los archivos comienzan (cabecera o header) con las letras 'BM' (0x42 0x4D), que lo identifica con el programa de visualización o edición. En la cabecera también se indica el tamaño de la imagen y con cuántos bytes se representa el color de cada píxel.

A continuación se detalla la estructura de la cabecera en la Tabla 8-4:

Tabla 8-4. Cabecera BMP.

Bytes	Información
0, 1	Tipo de fichero "BM"
2, 3, 4, 5	Tamaño del archivo
6, 7	Reservado
8, 9	Reservado
10, 11, 12, 13	Inicio de los datos de la imagen
14, 15, 16, 17	Tamaño de la cabecera del bitmap
18, 19, 20, 21	Anchura (píxels)
22, 23, 24, 25	Altura (píxels)
26, 27	Número de planos
28, 29	Tamaño de cada punto
30, 31, 32, 33	Compresión (0=no comprimido)
34, 35, 36, 37	Tamaño de la imagen
38, 39, 40, 41	Resolución horizontal
42, 43, 44, 45	Resolución vertical
46, 47, 48, 49	Tamaño de la tabla de color
50, 51, 52, 53	Contador de colores importantes

En la Figura 8-24 se muestra la estructura de este tipo de formato y en la Figura 8-25 la estructura de la cabecera.

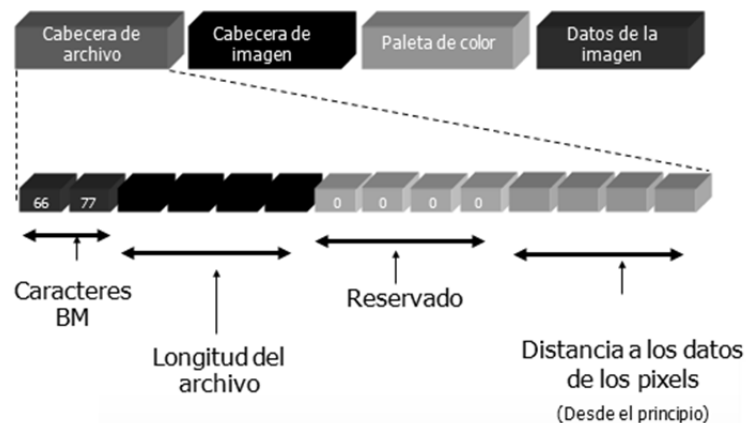


Figura 8-24. Estructura del archivo .bmp.

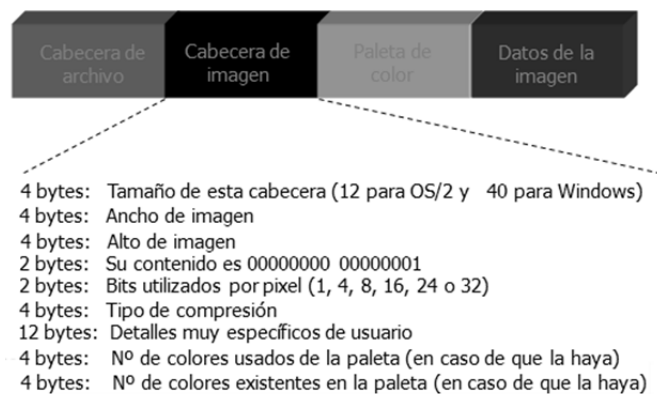


Figura 8-25. Estructura de la cabecera .bmp.

El tamaño de las imágenes es excesivamente grande, sobre todo en imágenes fotográficas. El tamaño viene dado por la ecuación (8-5)

$$\text{Tamaño de imagen (aprox.)} = \text{ancho} * \text{alto} * \text{bits_por_pixel. (8-5)}$$

No adecuado para transmisión por red. Poco popular fuera de los entornos de MS Windows (aunque está libre de patentes). Generalmente se usa en Aplicaciones que requieran una rápida salida por pantalla aplicaciones donde no deba haber pérdida de calidad, aun a costa del tamaño.

.TIF

TIFF (*Tagged Image File Format*) es un formato de archivo informático para imágenes. Las etiquetas también describen el tipo de compresión aplicado a cada imagen, que puede ser:

Sin compresión, Pack Bits, Compresión Huffman modificado, el mismo que las imágenes de fax (UIT grupo III y IV anteriormente CCITT), LZW, compresión sin pérdida y JPEG.

Hay también etiquetas que especifican el formato interno de almacenamiento de la imagen: completas, por bandas o por secciones rectangulares, lo cual permite a muchas aplicaciones optimizar los tiempos de carga o leer únicamente la zona de interés de una imagen grande.

Un aspecto muy práctico del formato TIFF es que permite almacenar más de una imagen en el mismo archivo.

Un mito que ha de desterrarse es la idea de que el formato TIFF no permite comprimir las imágenes. No obstante, algunas cámaras fotográficas digitales ofrecen la opción de grabar fotos en el formato TIFF, lo cual suele entenderse como sin compresión.

El formato TIFF admite opcionalmente el sistema de compresión sin pérdida de calidad, el conocido como LZW (Lempel-Ziv-Welch).

Es independiente de la plataforma, flexible y ampliable, puede adaptarse a muchos tipos de necesidades y puede contener (encapsular) ficheros con otros formatos. Pero es precisamente su flexibilidad lo que hace que sea difícil crear un programa adecuado para su manejo.

Se usa generalmente en edición fotográfica para impresiones de alta calidad. Aplicaciones con necesidades “especiales”, como imágenes multispectrales, con alta resolución de color, etc.

.GIF

(**CompuServe GIF**) es un formato gráfico utilizado ampliamente en la *World Wide Web*, tanto para imágenes como para animaciones.

El formato fue creado por CompuServe en 1987 para dotar de un formato de imagen en color para sus áreas de descarga de archivos, sustituyendo su temprano formato RLE en blanco y negro. GIF llegó a ser muy popular porque podía usar el algoritmo de compresión LZW (*Lempel Ziv Welch*) para realizar la compresión de la imagen, que era más eficiente que el algoritmo Run-length encoding (RLE) usado por los formatos PCX y MacPaint. Por lo tanto, imágenes de gran tamaño podían ser descargadas en un razonable periodo de tiempo, incluso con módems muy lentos.

GIF es un formato sin pérdida de calidad para imágenes con hasta 256 colores, limitados por una paleta restringida a este número de colores. Por ese motivo, con imágenes con más de 256 colores (profundidad de color superior a 8), la imagen debe adaptarse reduciendo sus colores, produciendo la consecuente pérdida de calidad. Una imagen GIF puede contener entre 2 y 256 colores (2, 4, 8, 16, 32, 64, 128 ó 256) entre 16,8 millones de su paleta. Por lo tanto, dado que la paleta tiene un número de colores limitado (no limitado en cuanto a colores diferentes), las imágenes que se obtenían con este formato por lo general eran muy pequeñas.

Sin embargo, dado que el algoritmo de compresión LZW estaba patentado, todos los editores de software que usaban imágenes GIF debían pagarle regalías a Unisys, la compañía propietaria de los derechos. Esta es una de las razones por las que el formato PNG se está volviendo cada vez más popular, en perjuicio del formato GIF.

Unisys, propietario de la patente del algoritmo LZW que se utiliza en el formato GIF reclamó durante años el pago de regalías por su uso. *Compuserve*, al desarrollar el formato, no sabía que el algoritmo LZW estaba cubierto por una patente. Debido a esto, cualquier programa capaz de abrir o guardar archivos GIF comprimidos con LZW debía cumplir con sus exigencias. Es necesario recalcar que el formato GIF puede utilizar otros métodos de compresión no cubiertos por patentes, como el método Run-length encoding.

Características no admitidas por otros formatos, como transparencias, animaciones y entrelazado, es adecuado para transmisión en redes. Por el contrario es poco adecuado para imágenes fotográficas: pérdida de color y tamaños muy grandes, por lo general se puede decir que es formato poco flexible.

Se usa en Compresión y almacenamiento de **dibujos e imágenes esquemáticas** con un número reducido de colores distintos. Transmisión de imágenes por **red**: imágenes de tamaño reducido (iconos, símbolos, etc.), animaciones sencillas.

-.jpg: (*Joint Photographic Experts Group*, Grupo Conjunto de Expertos en Fotografía) es el nombre de un comité de expertos que creó un estándar de compresión y codificación de archivos de imágenes fijas. Este comité fue integrado desde sus inicios por la fusión de varias agrupaciones en un intento de compartir y desarrollar su experiencia en la digitalización de imágenes. La ISO, tres años antes (abril de 1983), había iniciado sus investigaciones en el área.

Además de ser un método de compresión, es a menudo considerado como un formato de archivo. JPEG/Exif es el formato de imagen más común utilizado por las cámaras fotográficas digitales y otros dispositivos de captura de imagen, junto con JPG/JFIF, que también es otro formato para el almacenamiento y la transmisión de imágenes fotográficas en la World Wide Web. Estas variaciones de formatos a menudo no se distinguen, y se llaman JPEG.

En la mayoría de los casos, consigue un ratio compresión/calidad mucho mejor que otros formatos. Nivel de compresión ajustable generalmente entre 1:10 y 1:100. Es un formato muy popular y casi exclusivo en muchos ámbitos.

Como inconvenientes se puede destacar que los procesos de compresión/descompresión son complejos y costosos. E un formato que no incluye transparencias ni animaciones, genera “arteficios” (*artifacts*), y la información que se ha perdido en la obtención de la imagen con dicho formato no se recupera. Si trabajamos con un JPEG guardando en disco tras cada operación, la imagen se va degradando.

8.6 Cuestionario

1-La forma correcta de enfocar el haz del láser es:

- a) una circunferencia.
- b) Un círculo.
- c) Una elipse.
- d) Ninguna de las anteriores.

2-Régimen binario del DVD-Video es de:

- a) 9,6 Mbps.
- b) 26,16 Mbps.
- c) 10,08 Mbps.
- d) 6,1 Mbps.

3-¿De cuánto consta el bloque de un DVD en su estructura de datos?

- a) 2048
- b) 1024
- c) 2064
- d) Ninguna de las anteriores

4-¿Cuál es el Régimen binario del DVD-Audio?

- a) 6,1 Mbps.
- b) 9,6 Mbps.
- c) 4,3 Mbps.
- d) 26,16 Mbps.

5-El DVD-Video utiliza compresión tipo:

- a) MPEG I Layer 1.
- b) MPEG I Layer 2.
- c) MPEG 2.
- d) Ninguna de las anteriores.

6-La frecuencia de trama de subcódigo de un CD-DA es de:

- a) 59 Hz
- b) 60 Hz
- c) 75 Hz
- d) Ninguna de las anteriores

7-El DVD-18 tiene:

- a) Una cara y una capa de datos.
- b) Una cara y dos capas de datos.
- c) Dos caras y dos capas de datos.

d) Dos caras y una capa de datos.

8-Las Tarjetas SD:

a) Pueden soportar temperaturas de -45° a 85° .

b) Pueden soportar temperaturas de -25° a 85° .

c) Pueden soportar temperaturas de 0° a 25° .

d) Ninguna de las anteriores opciones es cierta.

9- La Capa reflectante del CD:

a) Separa el exterior del disco de la siguiente capa. El láser atraviesa el sustrato sin apenas variar su trayectoria.

b) Se modificadas durante la grabación pasando a contener la información grabada.

c) Permite que el rayo emitido se refleje en el disco, atraviere el sustrato y sus variaciones sean leídas por el sensor situado cerca del emisor.

d) Ninguna de las opciones anteriores es cierta.

10-El 7% de los datos de sobrecarga del CD son de control y visualización.

a) Verdadero.

b) Falso.

11-El número de bits de datos en una trama de un CD son:

a) 192 de 566.

b) 588 de 2024.

c) 12 de 32.

d) Ninguna de las opciones anteriores es cierta.

12-Si el bit1 del canal Q está a 1:

a) Se indica que la señal tienen cuatro canales.

b) Se indica que la señal tienen tres canales.

c) Se indica que la señal tienen dos canales.

d) Ninguna de las opciones anteriores es cierta.

13-Los fotodiodos receptores de los reproductores de CD:

a) Se estructuran de manera que cuatro son los de autoenfoco y dos de seguimiento de pista.

b) Son diferentes dependiendo de la información que contenga el CD.

c) Se estructuran de tal manera que dos son de autoenfoco y cuatro de seguimiento de pista.

d) Ninguna de las anteriores es cierta.

14-La longitud de onda del láser del Blu-Ray:

a) 780 nm.

b) 405 nm.

c) 650 nm.

d) Ninguna de las opciones anteriores es cierta.

15-La longitud de onda del láser del CD:

a) 780 nm.

b) 405 nm.

c) 650 nm.

d) Ninguna de las opciones anteriores es cierta.

16-La longitud de onda del láser del DVD:

a) 780 nm.

b) 405 nm.

c) 650 nm.

d) Ninguna de las opciones anteriores es cierta.

17-En cinta magnética únicamente se pueden almacenar señales analógicas.

a) Verdadero.

b) Falso.

18-Calcula el régimen binario de canal de una señal de 20 kHz en un disco duro. ($R_b=2 \cdot 20k=40k$ bps).

1-Códec viene de la abreviatura de codificador-decodificador.

a) Verdadero.

b) Falso.

19-Dentro de los formatos de compresión con pérdida podemos encontrar:

a) .wav, .aiff. y .wma.

b) .mp3 y .wma.

c) .wav o .aiff.

d) Ninguna de las anteriores es cierta.

20-La extensión .mp3 es usada en vídeos con codificación de vídeos mediante QuickTime.

a) Verdadero.

b) Falso.

21- El formato AAC:

a) Corresponde al estándar internacional "ISO/IEC 13818-7" como una extensión de MPEG-1.

b) Corresponde al estándar internacional "ISO/IEC 13818-7" como una extensión de MPEG-2.

c) Corresponde al estándar internacional "ISO/IEC 13818-7" como una extensión de MPEG-2-AAC.

d) Ninguna de las opciones anteriores es cierta.

22-Los controladores MIDI:

a) Generan los mensajes MIDI para activar o desactivar notas, variar un tono, etc.

b) Reciben los mensajes MIDI y los transforman en señales sonoras.

c) Aparatos destinados a grabar, reproducir o editar mensajes MIDI.

d) Ninguna de las opciones anteriores es correcta.

23-El conector usado en MIDI es:

- a) XLR.
- b) Jack.
- c) DIN-5.
- d) Coaxial.

9 CONCLUSIONES

La importancia y la gran cantidad de campos de aplicación de la imagen y el sonido hacen de este proyecto fuente de estudio muy útil para todo aquel que esté interesado en esta materia en continua expansión. No obstante la audiencia a la que se enfoca este documento es todo alumno que curse Sistemas Audiovisuales, lo cual hace que el libro tome forma de material didáctico más que de estudio científico o proyecto de investigación.

Para cada capítulo se establecen esquemas argumentales escalonados, es decir cada subapartado contiene conceptos que son de difícil comprensión si no se han estudiado los que le preceden. Es una forma hacer que la lectura sea fluida y establecer una guía para el alumno. Esto no obliga necesariamente a seguir esta guía, el lector más avanzado puede seleccionar sin problema el subapartado que le interese y entender lo que en éste se expone.

Una vez analizados todos los capítulos del proyecto y teniendo en cuenta que los objetivos principales del proyecto eran crear suficiente documentación de estudio para el alumno de Sistemas Audiovisuales que se imparte en la EUITT y asentar los conocimientos sobre imagen y sonido del autor se han llegado a las siguientes conclusiones:

- a) Al igual que se evalúa al docente al final del curso académico, para una correcta valoración de lo que en el documento se expone en necesaria una respuesta directa de todo aquel que la use. Es decir, aun que se han seguido unas pautas concretas de corrección y cada capítulo ha pasado el visto bueno del los profesores de la asignatura, se ha de preguntar a los lectores el grado de utilidad y conocimiento que da a este libro.
- b) Redactar una síntesis sobre esta materia requiere de tiempo y precisión. En principio se pensó en un periodo de tres meses de trabajo para realizarlo y sin embargo se han tardado prácticamente tres veces más de tiempo, ya que de otra forma se habrían cometido errores en la línea estructural y argumental o eliminado contenido importante, lo que al final habría tirado por tierra cualquier avance en el proyecto.
- c) La búsqueda de fuentes fiables es fundamental a la hora de realizar un estudio, compendio, libro o cualquier material de estudio técnico. Además es útil para poder redactar con la capacidad de incentivar al lector.

Este proyecto por tanto ayuda en el estudio de una materia que no sólo se aplica en tecnología, sino que como se explicó en un principio comporta cualidades artísticas, culturales y al fin y al cabo humanas. Luego a pesar de ser un documento técnico, la aplicación en la vida real de los conocimientos que aquí se desarrollan puede ser el mejorar la vida de otra persona.

..

10 BIBLIOGRAFÍA

- [1] Wikipédia, «TFT LCD,» [En línea]. Available:
http://es.wikipedia.org/wiki/TFT_LCD
[Último acceso: 06 12 2012].
- [2] Wikipédia, «Resolución en pantalla,» [En línea]. Available:
http://es.wikipedia.org/wiki/Resoluci%C3%B3n_de_pantalla
[Último acceso: 01 12 2012].
- [3] Wikipédia, «Proyector de vídeo,» [En línea]. Available:
http://es.wikipedia.org/wiki/Proyector_de_v%C3%ADdeo
[Último acceso: 01 12 2012].
- [4] Wikipédia, «Procesado digital de la luz,» [En línea]. Available:
http://es.wikipedia.org/wiki/Procesado_digital_de_luz
[Último acceso: 09 12 2012].
- [5] Wikipédia, «Pantalla LED,» [En línea]. Available:
http://es.wikipedia.org/wiki/Pantalla_LED
[Último acceso: 06 12 2012].
- [6] Wikipédia, «Pantalla láser,» [En línea]. Available:
http://es.wikipedia.org/wiki/Pantalla_l%C3%A1ser
[Último acceso: 07 12 2012].
- [7] Wikipédia, «Pantalla de Cristal Líquido,» [En línea]. Available:
http://es.wikipedia.org/wiki/Pantalla_de_cristal_l%C3%ADquido
[Último acceso: 3 12 2012].
- [8] Wikipédia, «Panel SED,» Panel SED, [En línea]. Available:
http://es.wikipedia.org/wiki/Panel_SED
[Último acceso: 07 12 2012].
- [9] Wikipédia, «OLED,» [En línea]. Available:
http://es.wikipedia.org/wiki/Diodo_org%C3%A1nico_de_emisi%C3%B3n_de_luz
[Último acceso: 06 12 2012].
- [10] Wikipédia, «Monitores de computadora,» [En línea]. Available:
http://es.wikipedia.org/wiki/Monitor_de_computadora
[Último acceso: 3 12 2012].

- [11] Wikipédia, «Monitor de vídeo,» [En línea]. Available:
http://es.wikipedia.org/wiki/Monitor_de_v%C3%ADdeo
[Último acceso: 1 12 2012].
- [12] Wikipédia, «Cristal líquido sobre silicio,» [En línea]. Available:
http://es.wikipedia.org/wiki/Cristal_l%C3%ADquido_sobre_silicio
[Último acceso: 08 12 2012].
- [13] J. C. Whitaker y B. Benson, *Standard Handbook of Video and Television Engineering*, NY, USA: McGraw-Hill, 2003.
- [14] C. P. Vega, «Tubo de rayos catódicos,» [En línea]. Available:
<http://personales.unican.es/perezvr/pdf/TUBOS%20DE%20RAYOS%20CATODICOS.pdf>
[Último acceso: 3 12 2012].
- [15] F. Rumsey, *Reproduction and recording audio*, USA: Elsevier, 2004 .
- [16] K. C. Pohlmann, *Principles of Digital Audio*, USA: McGraw-Hill Companies Incorporated, 2005.
- [17] L. o. Plasma, «Cómo Funciona y qué es el contraste dinámico,» [En línea]. Available:
<http://www.lcdoplasma.es/como-funciona-y-que-es-el-contraste-dinamico/>
[Último acceso: 01 12 2012].
- [18] L. I. Ortiz Berenguer y J. L. Rodríguez Vásquez, *Ingeniería de Vídeo en entornos digitales*, Madrid, España: Dpto. Publicaciones EUITT, 2008.
- [19] A. Martín Marcos, *Televisión Digital Vol.I. Compresión MPEG*, Madrid, España: Ciencia 3, 2006.
- [20] A. Martín Marcos, *Sistemas de Televisión*, Madrid, España: Editorial Ciencia 3, 1996.
- [21] A. Martín Marcos y P. Martín Martín, *Colorimetría*, Madrid, España: Ciencia, 2000.
- [22] A. C. Luther, *Principles of Digital Audio and Video*, MA, USA: Artech House, Inc., 1997.
- [23] P. labs, «OLED bileşenleri,» [En línea]. Available:
<http://www.pclabs.com.tr/2008/10/27/oled-nedir-nasil-calisir/>
[Último acceso: 06 12 2012]
- [24] ITU, «Recommendation BT.656,» 1986-1992-1994-1995-1998.
- [25] ITU, «Recommendation BT.601,» 1982.
- [26] ISO/IEC, «11172-1, Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s » 1993.

- [27] B. D. d. ILCE, «Cristales Líquidos,» [En línea]. Available:
http://bibliotecadigital.ilce.edu.mx/sites/ciencia/volumen2/ciencia3/104/htm/sec_6.htm
[Último acceso: 4 12 2012].
- [28] J. González Rodríguez y A. Mínguez Olivar, *Procesado digital de la señal audiovisual*, Madrid, España: Dpto. Publicaciones EUITT, 2001.
- [29] E. Q. Gómez, «El monitor: tipos y características,» [En línea]. Available:
<http://recursostic.educacion.es/observatorio/web/es/equipamiento-tecnologico/hardware/268-eduardo-quiroga-gomez>
[Último acceso: 01 12 2012].
- [30] G. Davis, *Sound Reinforcement Handbook*, USA: Hal Leonard Corporation, 1987.
- [31] D. O. Carrasco, «Tipo de conexiones multimedia,» [En línea]. Available:
<http://recursostic.educacion.es/observatorio/web/es/equipamiento-tecnologico/hardware/1002-tipos-de-conexiones-multimedia>
[Último acceso: 1 12 2012].
- [32] J. L. S. Bote, *Micrófonos*, Madrid, España: Dpto. Publicaciones EUITT, 2002.
- [33] J. L. S. Bote, *Altavoces*, filtros de cruce y bocinas, Madrid, España: Dpto. Publicaciones EUITT, 2002.

ANEXO A

Dithering

La correlación entre el error de cuantificación y la señal es muy pequeña en señales de alto nivel, lo que lo hace aleatorio y que sea percibido como ruido blanco. En señales de bajo nivel ocurre todo lo contrario, el error de cuantificación depende de la amplitud de la señal de entrada provocando la distorsión de la misma.

El Dither es una técnica digital usada para eliminar los problemas causados por el error de cuantificación en señales de bajo nivel. Consiste en la suma, antes del muestreo, de ruido aleatorio sin relación alguna con la señal a digitalizar para alinear el proceso de cuantificación. Con Dither la señal de audio cambia de valores constantemente respecto a los niveles de cuantificación. En vez de que se produzcan periódicamente los patrones de cuantificación, cada ciclo de la onda es diferente, lo que hace que el error de cuantificación sea aleatorio y no dependa de la señal. De esta manera se elimina el problema de la distorsión en niveles bajos, pero se añade algo de ruido a la salida de la señal de audio que puede ser percibido.

El Dither no enmascara el error de cuantificación, sino que permite que el sistema digital codifique amplitudes más pequeñas que el bit menos significativo (LSB). Un sistema correctamente diseñado con un Dither supera a un sistema analógico en su relación S/N. En cambio si no se tuviese el Dither, el sistema digital sería peor, sobre todo con señales de nivel bajo. En la Figura 0-1 se puede ver el efecto del Dither.

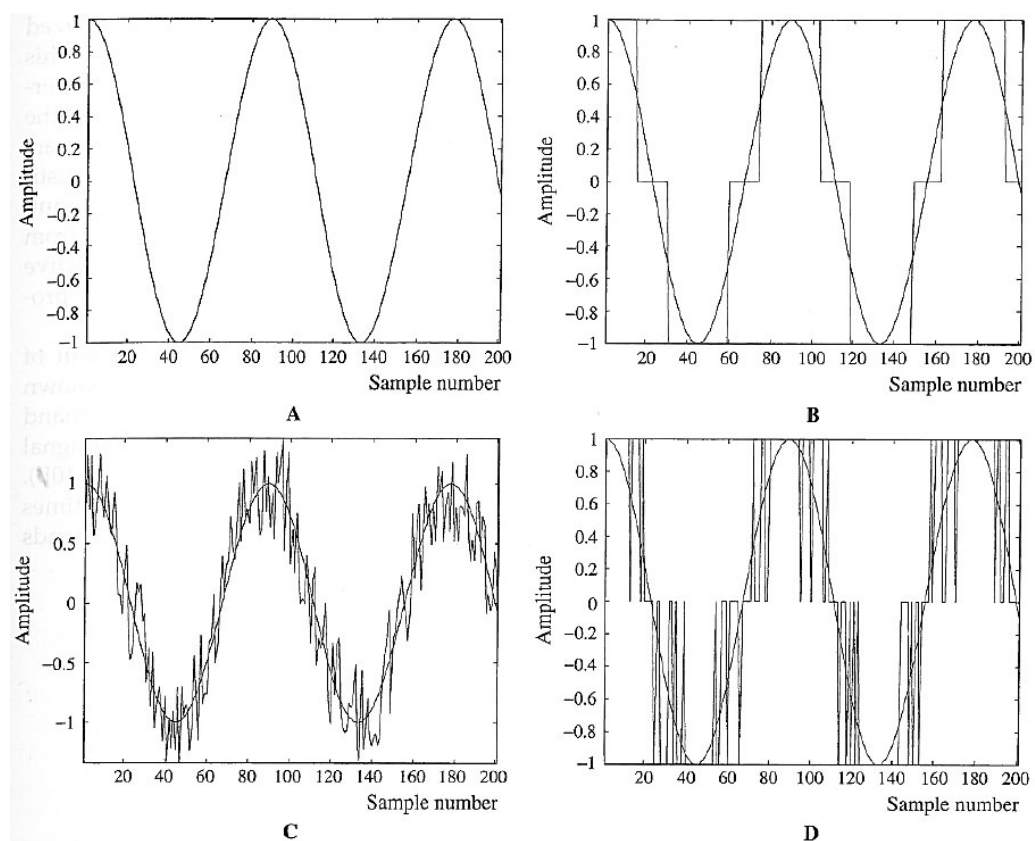


Figura 0-1. Dither usado para aliviar los efectos del error de cuantificación. A: Señal sin Dither de 2 LSB de amplitud. B: Resultados de cuantificación en una cuantificación de 3 niveles. C: Señal con Dither. D: Resultados de cuantificación en una cuantificación de 3 niveles con Dither.

Considere una señal de audio con una amplitud en torno a un intervalo de cuantificación. La señal de entrada podría estar comprendida completamente dentro del intervalo de cuantificación, lo que resultaría una señal cuantificada de valor constante (continua), o podría por el contrario estar variando alrededor del umbral de

cuantificación produciendo una señal cuadrada como se muestra en la Figura 0-1A-B. Esto demuestra que el cuantificador, con niveles bajos de entrada puede actuar como un limitador de nivel; en otras palabras, aparece una gran distorsión. El efecto es completamente diferente cuando se usa Dither. El resultado se muestra en la Figura 0-1C-D, es una señal pulsante que preserva la información de la señal de audio. La señal cuantificada conmuta arriba y abajo, de acuerdo con la variación que proporciona la señal de Dither, siguiendo el valor medio de la señal de entrada. La información está codificada en la variación de la anchura de los pulsos de la señal cuantificada. Este tipo de codificación se denomina modulación por anchura de pulsos, y preserva de forma precisa la forma de onda de la señal de entrada. El valor medio de la señal cuantificada oscila continuamente entre dos niveles, mitigando los efectos del error de cuantificación. De forma similar, el ruido analógico podría codificarse con una señal binaria; los valores 0 y 1 aparecerían en el bit menos significativo en cada instante de muestreo, manteniendo en el espectro el carácter de ruido blanco. El resultado perceptual: señal original sumada con ruido, un resultado más deseable que la anterior onda cuadrada.

Cálculo de la Relación señal/ruido

Para caracterizar los dispositivos digitales se puede determinar la relación entre el valor máximo de la señal y el valor máximo del error de cuantificación: la relación señal-error (S/E).

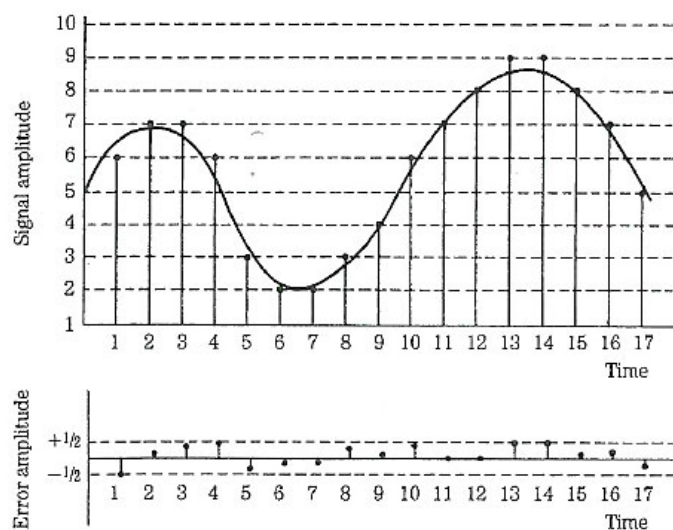


Figura 0-2. El valor de amplitud es redondeado al escalón de cuantificación más cercano. El error de cuantificación en un instante de muestreo es menor o igual a $\frac{1}{2} Q$.

Se considera un sistema de cuantificación que trabaja con n bits, N es el número de escalones de cuantificación viene dado por la expresión (0-1):

$$N = 2^n. \quad (0-1)$$

La mitad de esos N niveles son usados para codificar la parte positiva y negativa de la forma de onda. Si Q es el intervalo de cuantificación, los valores de pico de la señal son $\pm Q2^{n-1}$. Si se asume que la señal de entrada es sinusoidal, el promedio máximo de la señal (RMS) se calcula con la expresión (0-2):

$$S_{RMS} = \frac{Q2^{n-1}}{\sqrt{2}}. \quad (0-2)$$

La energía del error de cuantificación también se puede determinar. Cuando la señal de entrada tiene nivel alto y un espectro amplio, el error de cuantificación es estáticamente independiente y se distribuye de manera uniforme entre $+Q/2$ y $-Q/2$, incluyendo el valor nulo, donde Q es el intervalo de cuantificación. Esto establece una función de densidad de probabilidad uniforme con amplitud de $1/Q$. El error es aleatorio de muestra a muestra y

su espectro es plano. Ignorando el error fuera de la banda de la señal, el error de cuantificación RMS, E_{RMS} , se puede hallar integrando el producto del error, e^2 , por la probabilidad de este, $p(e) = 1/Q$, tal y como se muestra en el desarrollo de la expresión (0-3):

$$E_{RMS} = \sqrt{\int_{-\infty}^{+\infty} e^2 p(e) de} = (0-3)$$

$$= \sqrt{\frac{1}{Q} \int_{-Q/2}^{+Q/2} e^2 de} = \sqrt{\frac{Q^2}{12}} = \frac{Q}{\sqrt{12}}$$

Se puede establecer así la relación señal-error como la que de la expresión (0-4):

$$\frac{S}{E} = \left[\frac{S_{RMS}}{E_{RMS}} \right]^2 = \frac{\left[\frac{Q 2^{n-1}}{\sqrt{2}} \right]^2}{\left[\frac{Q}{\sqrt{12}} \right]^2} = \frac{3}{2} (2^{2n}). (0-4)$$

En decibelios se calcula por la expresión (0-5):

$$\begin{aligned} \frac{S}{E} (dB) &= 10 \log \left[\frac{3}{2} (2^{2n}) \right] \\ &= 20 \log \left[\sqrt{\frac{3}{2}} 2^n \right] = 6,02 \cdot n + 1,76 \text{ dB. (0-5)} \end{aligned}$$

Usando esta aproximación se observa que cada vez que se añade un bit a la palabra digital de audio, la relación S/E aumenta 6 dB, o lo que es lo mismo, se multiplica por dos. Por ejemplo, un sistema que cuente con 15 bits tendrá una relación S/E = 92 dB, mientras que si se tuviesen 16 bits la relación sería de 98 dB. Es importante decir que este resultado se asume para señales de nivel alto, con las que el error de cuantificación se reparte de forma uniforme por el espectro audible percibiéndose como un ruido blanco. Para el caso de señales de nivel bajo la distorsión puede hacerse notable.

ANEXO B

Obtención de los códigos VLC para los bloques de luma y croma

Para establecer la referencia para los bloques de luminancia, el bloque 0 se codifica respecto al 3 del macrobloque anterior. El bloque 1 se referencia al bloque 0, el bloque 2 se referencia al 1, el bloque 3 se referencia al 2 y a su vez este bloque 3 da origen al predictor para el bloque 0 del siguiente macrobloque. Para los bloques de crominancia su referencia se establece en su homólogo del macrobloque anterior.

Si el formato de muestreo fuese otro cualquiera diferente del 4:2:0, siempre se codificará respecto de su homólogo anterior, que en estos casos aun siendo de crominancia podrá ser del mismo macrobloque o del anterior.

Al comienzo de cada slice, los predictores para los tres tipos de bloques (Y, Cr, y Cb), se resetean a los valores que se encuentran expresados en la columna de la derecha de la tabla x, que es el valor central de la indeterminación existente, como no podía ser de otra forma.

A partir de ese momento y hasta que acabe el slice, la predicción se hace respecto al bloque anterior, tal como se acaba de describir, pero para el primer bloque no existe otro anterior y se toma el que se ha establecido en la referida tabla. También se tomará ese valor central si el macrobloque no es Intra o cuando se haya saltado el anterior macrobloque (*skipped macroblock*), indicando por la situación de que el incremento de dirección del macrobloque sea superior a 1.

Para codificar el valor diferencial, se establece un código formado por dos variables. La primera es un código VLC (código de longitud variable) que indica el tamaño de la segunda. La primera parte se denomina tamaño y se codifica diferente de el bloque es de luminancia o de crominancia, siendo los códigos VLC que especifican el tamaño, los que se representan en las Tablas 0-1 :

Tablas 0-1. Códigos VLC de luma y croma.

Código VLC	dct_dct_size_luminance	Código VLC	dct_dc_size_chrominance
100	0	00	0
00	1	01	1
01	2	10	2
101	3	110	3
110	4	1110	4
1110	5	11110	5
11110	6	111110	6
111110	7	1111110	7
1111110	8	11111110	8
11111110	9	111111110	9
111111110	10	1111111110	10
111111111	11	1111111111	11

Por ejemplo, si un bloque de luminancia aparece el código 111110, observando la tabla, eso quiere decir al decodificador que los próximos 7 bits que van a llegar a continuación de esta primera parte indican el valor del índice diferencial, que puede ser positivo o negativo, nunca será cero. A valores positivos se les codificará directamente su valor en binario y a los negativos se expresa el módulo en binario natural y luego se cambian ceros por unos y unos por ceros.

A continuación se explicará todo lo anterior con un ejemplo

Bloque de luminancia:

Valor anterior: 45.

Valor actual: 62.

Escalón cuántico de trabajo: 1 (intra_dc_precision= '1')

Valor DC a enviar diferencialmente $62 - 45 = +17$

Como es escalón unidad: $QDC = DC//1 = +17//1 = +17$

Este valor necesita 5 dígitos para ser expresado (10001 en binario). El código VLC será 1110.

Nos quedaría por tanto un código para este coeficiente de: 1110 10001 (9 bits en total).

Codificación de los índices AC tipo Intra

Después de esta exploración en zig-zag se dispone de una secuencia unidimensional que tiene la siguiente estructura:

$$I_0, I_1, I_2, I_3, I_4, \dots, I_{63}$$

Dadas las características de la matriz de índices, al final de la secuencia completa de 62 índices AC aparecerán muchos valores nulos. Obviamente esos valores no serán codificados, sino que cuando en la secuencia de 63 índices aparezca el último valor significativo (distinto de cero), ya no se siguen codificando los ceros que vienen detrás, sino que se coloca un código conocido como Fin de Bloque, que será interpretado precisamente como eso, como todos los índices que aún no se han decodificado pasan a tomar inmediatamente el valor cero. La situación será la siguiente:

$$I_1, I_2, I_3, I_4, \dots, I_{\text{FINAL}}$$

En la mayoría de los casos el número de valores obtenidos para el total de la secuencia no será muy alto, ya que la ventaja de la DCT y la posterior cuantificación es que la mayoría de los índices en las proximidades de la esquina inferior derecha son nulos. Estando inmersos en la codificación de macrobloques tipo I, a partir de la secuencia unidimensional de índices obtenida en el apartado anterior, es necesario proceder a su codificación.

Como bastantes de los índices intermedios pueden también ser ceros, para codificar la secuencia se utiliza la técnica conocida como *run and level* ya utilizada en la codificación JPEG, que consiste en la creación de un macrosímbolo formado por todos los ceros que se encuentran antes del valor 'no cero' y seguidos justamente de ese valor 'no cero'. Ese macrosímbolo es el que se codifica en conjunto. Para su codificación se acude a una Tabla en la que mediante un método Huffman y en función de las probabilidades de aparición se han establecido los códigos correspondientes, de manera, lógicamente, que ninguno de ellos es prefijo de otro, para que no haya ninguna ambigüedad en el proceso de decodificación.

En realidad hay dos tablas disponibles, debiendo de utilizarse una u otra en función del tipo de bloque a codificar y de lo que indiquen una serie de variables. La tabla más utilizada es la denominada en la Norma B.14, pero en algunos casos deberá utilizarse la tabla denominada B.15. La utilización de una u otra viene condicionada por dos variables:

- La variable *intra_vic_format* que con un solo bit se aloja en la cabecera de la extensión para la codificación de la imagen.
- La variable *macroblock_intra* que se encuentra alojada a su vez en la variable *macroblock_type* que es más amplia. La variable a la que nos referimos también tiene un solo bit.

Siempre se utiliza la tabla B14, salvo en el caso de que las dos variables anteriores sean '1' (que sea un bloque codificado tipo 'intra' y que se active la variable *intra_vic_format*) en cuyo caso se utilizará la tabla B15.

En la Tabla 0-2. Tipo de tabla a usar en función de las variables *intra_vic_format* y *macroblock_intra* se representa esto que se acaba de explicar.

Tabla 0-2. Tipo de tabla a usar en función de las variables *intra_vic_format* y *macroblock_intra*

<i>intra_vic_format</i>	0	1
Bloque tipo I (<i>macroblock_intra</i> = 1)	Tabla B.14	Tabla B.15
Bloque tipo no I (<i>macroblock_intra</i> = 0)	Tabla B.14	Tabla B.14

El proceso seguido por el codificador es muy simple y básicamente se resume en las siguientes acciones:

-Se mira el número de ceros que hay antes de un valor que no sea cero. A ese número de ceros es a lo que se denomina 'run' o abreviadamente RR. Al valor distinto de cero que se encuentra al final es lo que se denomina 'level'.

-Con estos dos valores se va a la tabla (normalmente la) y se saca el código correspondiente.

Por ejemplo, si la secuencia a codificar fuese ...0002...esto quiere decir que hay que mirar en la columna en la que $run = 3$ y $level = 2$. Yendo a la tabla se localiza el código que le corresponde, que en este caso es 00100100s. Si el valor es positivo (como es este caso en el que vale +2) el bit 's' será cero, pero si hubiese sido negativo ese último bit debería de haber sido un 1.

Código para 0002 → 001001000

Código para 0002 → 001001001

Cuando se utilice la tabla B.14 se utilizará el código binario '10' (conocido como EOB o *End of Block*) para indicar que se intenta comunicar el concepto de *Fin de bloque* y que todos los índices significativos ya han sido recibidos, debiendo considerar en ese bloque ceros a todos los índices que aún no se han recibido. Como se indica al final de la propia tabla, este código no puede estar como único código del bloque, sino que con anterioridad se habrán recibido otros códigos.

Tabla 0-3. Tabla B.14

Código de longitud variable (Nota 1)	Pasada (run)	Nivel (level)
10 (Nota 2)	Fin de bloque	
1 s (Nota 3)	0	1
11 s (Nota 4)	0	1
011 s	1	1
0100 s	0	2
0101 s	2	1
0010 1 s	0	3
0011 1 s	3	1
0011 0 s	4	1
0001 10 s	1	2
0001 11 s	5	1
0001 01 s	6	1
0001 00 s	7	1
0000 110 s	0	4
0000 100 s	2	2
0000 111 s	8	1
0000 101 s	9	1
0000 01	Escape	
0010 0110 s	0	5
0010 0001 s	0	6
0010 0101 s	1	3
0010 0100 s	3	2
0010 0111 s	10	1
0010 0011 s	11	1
0010 0010 s	12	1
0010 0000 s	13	1
0000 0010 10 s	0	7
0000 0011 00 s	1	4
0000 0010 11 s	2	3
0000 0011 11 s	4	2
0000 0010 01 s	5	2
0000 0011 10 s	14	1
0000 0011 01 s	15	1
0000 0010 00 s	16	1
0000 0001 1101 s	0	8
0000 0001 1000 s	0	9
0000 0001 0011 s	0	10
0000 0001 0000 s	0	11
0000 0001 1011 s	1	5

Código de longitud variable (Nota 1)	Pasada (run)	Nivel (level)
0000 0001 0100 s	2	4
0000 0001 1100 s	3	3
0000 0001 0010 s	4	3
0000 0001 1110 s	6	2
0000 0001 0101 s	7	2
0000 0001 0001 s	8	2
0000 0001 1111 s	17	1
0000 0001 1010 s	18	1
0000 0001 1001 s	19	1
0000 0001 0111 s	20	1
0000 0001 0110 s	21	1
0000 0000 1101 0 s	0	12
0000 0000 1100 1 s	0	13
0000 0000 1100 0 s	0	14
0000 0000 1011 1 s	0	15
0000 0000 1011 0 s	1	6
0000 0000 1010 1 s	1	7
0000 0000 1010 0 s	2	5
0000 0000 1001 1 s	3	4
0000 0000 1001 0 s	5	3
0000 0000 1000 1 s	9	2
0000 0000 1000 0 s	10	2
0000 0000 1111 1 s	22	1
0000 0000 1111 0 s	23	1
0000 0000 1110 1 s	24	1
0000 0000 1110 0 s	25	1
0000 0000 1101 1 s	26	1
0000 0000 0111 11 s	0	16
0000 0000 0111 10 s	0	17
0000 0000 0111 01 s	0	18
0000 0000 0111 00 s	0	19
0000 0000 0110 11 s	0	20
0000 0000 0110 10 s	0	21
0000 0000 0110 01 s	0	22
0000 0000 0110 00 s	0	23
0000 0000 0101 11 s	0	24
0000 0000 0101 10 s	0	25
0000 0000 0101 01 s	0	26
0000 0000 0101 00 s	0	27

Código de longitud variable (Nota 1)	Pasada (run)	Nivel (level)
0000 0000 0100 11 s	0	28
0000 0000 0100 10 s	0	29
0000 0000 0100 01 s	0	30
0000 0000 0100 00 s	0	31
0000 0000 0011 000 s	0	32
0000 0000 0010 111 s	0	33
0000 0000 0010 110 s	0	34
0000 0000 0010 101 s	0	35
0000 0000 0010 100 s	0	36
0000 0000 0010 011 s	0	37
0000 0000 0010 010 s	0	38
0000 0000 0010 001 s	0	39
0000 0000 0010 000 s	0	40
0000 0000 0011 111 s	1	8
0000 0000 0011 110 s	1	9
0000 0000 0011 101 s	1	10
0000 0000 0011 100 s	1	11
0000 0000 0011 011 s	1	12
0000 0000 0011 010 s	1	13
0000 0000 0011 001 s	1	14
0000 0000 0001 0011 s	1	15
0000 0000 0001 0010 s	1	16
0000 0000 0001 0001 s	1	17
0000 0000 0001 0000 s	1	1
0000 0000 0001 0100 s	6	3
0000 0000 0001 1010 s	11	2
0000 0000 0001 1001 s	12	2
0000 0000 0001 1000 s	13	2
0000 0000 0001 0111 s	14	2
0000 0000 0001 0110 s	15	2
0000 0000 0001 0101 s	16	2
0000 0000 0001 1111 s	27	1
0000 0000 0001 1110 s	28	1
0000 0000 0001 1101 s	29	1
0000 0000 0001 1100 s	30	1
0000 0000 0001 1011 s	31	1

NOTA

- 1 El último bit 's' da el signo del nivel: '0' para positivo, '1' para negativo.
- 2 «Fin de bloque» no será el único código del bloque.
- 3 Este código se utilizará para el primer coeficiente (DC) en el bloque.
- 4 Este código se utilizará para todos los otros coeficientes.

Puede que la secuencia a codificar no esté presente en la tabla anterior, para lo cual se guarda un código denominado ESCAPE que es el 000001 y que lógicamente no puede confundirse con ningún otro. Cuando aparece ESCAPE, los siguientes 6 bits indican el número de ceros antes de un 1 y los 12 siguientes indican el valor del índice en sí. En la Tabla 0-4 se muestra un ejemplo de un código escape.

Tabla 0-4. Ejemplo de un código escape

ESCAPE	Número de ceros	Valor del índice
0000001	000011 (3)	100000000010 (-2046)